

**Exploring  
regulatory functions and  
enzymatic activities  
in the nidovirus replicase**

**Danny Nedialkova |**



# Exploring regulatory functions and enzymatic activities in the nidovirus replicase

**Proefschrift**

ter verkrijging van  
de graad van Doctor aan de Universiteit Leiden,  
op gezag van Rector Magnificus prof.mr. P.F. van der Heijden,  
volgens besluit van het College voor Promoties  
te verdedigen op woensdag 23 juni 2010  
klokke 11.15 uur

door

**Danny Doncheva Nedialkova**

geboren te Sofia, Bulgarije  
in 1983

## **Promotiecomissie**

**Promotores:** Prof. Dr. E.J. Snijder  
Prof. Dr. A.E. Gorbalenya

**Overige leden:** Prof. Dr. W. J. M. Spaan  
Prof. Dr. N. Tautz (University of Lübeck)  
Dr. R.J. de Groot (Universiteit Utrecht)

The research described in this thesis was carried out at the Department of Medical Microbiology of the Leiden University Medical Center, Leiden, The Netherlands, and was financially supported by grant 700.52.306 from the Council for Chemical Sciences of the Netherlands Organization for Scientific Research (NWO-CW).

*"Science is not about building a body of known 'facts'.  
It is a method for asking awkward questions and subjecting them to a reality-check,  
thus avoiding the human tendency to believe whatever makes us feel good."*

Terry Pratchett



## CONTENTS

CHAPTER 1	General introduction	9
CHAPTER 2	Regulation of the plus-strand RNA virus replicative cycle:   putting it together by taking it apart	31
CHAPTER 3	Site-directed mutagenesis of the nidovirus replicative   endoribonuclease NendoU exerts pleiotropic effects on   the arterivirus life cycle	81
CHAPTER 4	Biochemical characterization of arterivirus nonstructural   protein 11 reveals the nidovirus-wide conservation of a   replicative endoribonuclease	101
CHAPTER 5	Arterivirus subgenomic mRNA synthesis and virion   biogenesis depend on the multifunctional nsp1   autoprotease	127
CHAPTER 6	Arterivirus nsp1 modulates the accumulation of minus-   strand templates to control the relative abundance of viral   mRNAs	149
CHAPTER 7	Expression and purification of recombinant equine   arteritis virus nonstructural protein 1	179
CHAPTER 8	General discussion	203
SUMMARY		229
SAMENVATTING		231
CURRICULUM VITAE		237





# Chapter 1

**General introduction |**



## INFECTIOUS AGENTS WITH RNA GENOMES

Ribonucleic acid (RNA) is a conformationally flexible biopolymer composed of nucleotide subunits. Each nucleotide consists of a phosphate group and a pentose sugar (ribose), together forming the backbone of every RNA chain, and one of four nitrogenous bases (adenine, guanine, cytosine or uracil). RNA molecules are characterized by a high degree of functional plasticity and are critically involved in fundamental cellular processes, including the priming of DNA replication, the transfer of genetic information to the translation apparatus, and the posttranscriptional regulation of gene expression. RNA can also have catalytic properties. Chemical transformations that can be catalyzed by RNA molecules include RNA-processing events, nucleotide synthesis, RNA polymerization and peptide bond formation<sup>3,13,47,54,62,103,108</sup>. The remarkable ability of RNA to serve as both a carrier of genetic information and a catalyst for the template-based duplication of this information lends credence to the hypothesis that the first self-replicating systems that emerged on Earth may have been based on RNA only<sup>37,56</sup>. The conformational variability and chemical functionality of RNA is restricted by the fact that RNA polymers are composed of just four building blocks with limited chemically reactive groups. These properties of RNA, combined with the necessity to respond to arising environmental challenges, may have led to the invention of protein synthesis templated by RNA sequences. The transfer of catalytic “responsibilities” to this novel and more versatile biomolecule marked the transition to protein-based metabolism. Subsequently, DNA replaced RNA in its role of genetic material due to its greater chemical stability, permitting much larger genomes. Today, the only known entities that rely on RNA for the storage of their genetic information are RNA viruses, viroids and virus-associated RNAs, suggesting they may be “evolutionary relics of the RNA world”.

Phytopathogenic viroids are circular, non-coding RNAs with lengths between 250 and 400 nucleotides (nt) that are not encased in a protein shell, and their entire replicative cycle depends on host factors. Viroid RNAs are capable of utilizing cellular DNA-dependent RNA polymerases (DdRp) for their replication, and spread through infected plants via plasmodesmata (reviewed in reference<sup>102</sup>). A subviral human pathogen, hepatitis delta virus (HDV), shares the ability of viroids to use host cell DdRp for amplifying its circular, single-stranded RNA genome of 1.6 kb. In contrast to viroids, HDV encodes a protein required for replication of its RNA – the HD antigen (HDAg) is expressed from the antigenomic transcript and encapsidates HDV genomes. The envelope proteins on the outer surface of HDV are entirely provided by hepatitis B virus (HBV), making HDV dependent on host cells for RNA replication and on another virus for particle assembly<sup>105</sup>.

Viruses that replicate only via RNA intermediates encode a viral RNA-dependent RNA polymerase (RdRp) to amplify their genomes. Plus-strand RNA (+RNA) viruses carry linear, single-stranded genomes of mRNA polarity that can initiate viral gene expression immediately after genome uncoating in infected cells. The genomes of minus-strand RNA (-RNA) viruses, by contrast, are complementary to mRNA and need to be transcribed into mRNAs before viral genes can be expressed. To this end, -RNA virus particles also

deliver a viral “transcriptase” into their host cells. The single-stranded RNA genomes of retroviruses, although of positive polarity, also do not function as mRNAs at the onset of infection, but are transcribed into DNA intermediates by a virion-associated RNA-dependent DNA polymerase.

Although the length of RNA genomes is presumably constrained by the low fidelity of their RNA synthesis<sup>31,53</sup>, +RNA viruses constitute the largest virus group today and comprise ubiquitous pathogens of plants and animals. In their most economic form, +RNA genomes resemble cellular mRNAs, encoding a single open reading frame (ORF), flanked by 5′ and 3′ untranslated sequences. The latter carry signals which ensure the recognition and efficient translation of the genomic RNA by the cellular translation apparatus, as well as its replication by the viral RNA-synthesizing complex. Polypeptides expressed from the viral ORF invariably include an RNA-dependent RNA polymerase (RdRp), the enzyme responsible for RNA-templated RNA synthesis, a process which takes place exclusively in the cytoplasm of the infected cell. In order to express multiple viral gene products from a single ORF, many +RNA viruses rely on proteolytic processing of polypeptide precursors into mature protein subunits. In addition, eukaryotic +RNA viruses with polycistronic genomes have evolved different ways to encode and express multiple ORFs. Many plant and insect +RNA viruses employ bi-segmented and occasionally tri-segmented genomes. Several translation regulation strategies, such as “leaky scanning” by ribosomes and ribosomal frameshifting, can ensure the translation of adjacent or overlapping viral ORFs from the same mRNA molecule. Expression of ORFs positioned internally on the genomic RNA can involve the use of internal ribosome entry sites (IRES) for ribosome recruitment or the synthesis of subgenomic (sg) mRNAs in which these ORFs occupy a 5′-proximal position. These strategies are not mutually exclusive, and +RNA viruses may utilize several of these mechanisms for the expression of their genetic information.

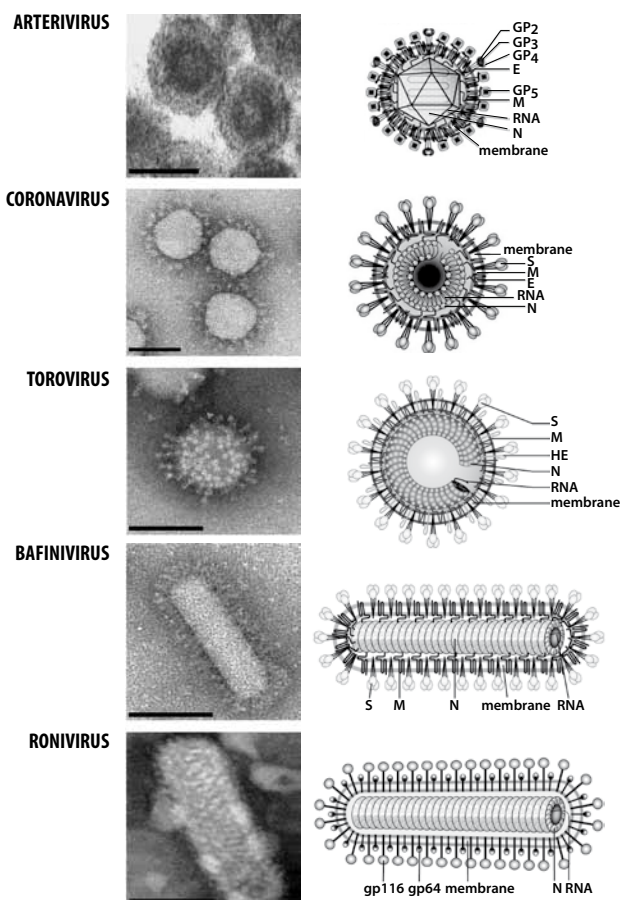
In addition to the viral RdRp, +RNA viruses generally encode multiple protein factors that support RdRp activity and/or are involved in the spatial and temporal regulation of viral RNA synthesis. The presence of one such RdRp co-factor - an RNA helicase, among the replicative proteins of +RNA viruses was postulated to have permitted genomic RNA expansion to sizes above ~6 kb. This hypothesis is based on the selective absence of this domain in +RNA viruses with genome sizes below ~6 kb, and the presumed role of viral helicases in unwinding large double-stranded (ds) RNA regions of replicative intermediates, as well as structured regions present in template RNAs that may hamper RdRp processivity<sup>39,40</sup>. The greater coding capacity, in turn, may have permitted the evolution of viral gene products solely dedicated to modulating and evading antiviral host responses, allowing for greater plasticity of these viral genomes and increased ability of the virus to adapt to different host environments.

## NIDOVIRUSES

Three distantly related families of enveloped +RNA viruses – *Coronaviridae*, *Arteriviridae* and *Roniviridae*, have been united in the order of *Nidovirales* based on similarities in their gene organization, expression strategies, and the presumed common ancestry of their key replicative enzymes<sup>17,38,39</sup>. The *Coronaviridae* family has recently been reorganized and now comprises the two subfamilies of *Coronavirinae* (subdivided into the *Alpha-*, *Beta-*, and *Gammacoronavirus* genera) and *Torovirinae* (genera *Bafinivirus* and *Torovirus*) (<http://www.ictvonline.org/virusTaxonomy.asp?version=2009>).

Nidoviruses infect a wide variety of host, ranging from invertebrates to mammals. The ronivirus yellow head virus (YHV) causes significant mortality of farmed prawns, resulting in severe economic damage through heavy production losses<sup>34</sup>. The *Arteriviridae* family also includes economically important animal pathogens, such as the porcine reproductive and respiratory syndrome virus (PRRSV), which is associated with significant economic losses in the swine industry<sup>71</sup>. Members of *Coronaviridae* comprise numerous enteric and/or respiratory pathogens of farm and companion animals, and the *Coronaviridae* family is currently the only major nidovirus branch that includes known human pathogens. After rhinoviruses, human coronaviruses are the second most common cause of common cold-like disease in humans. In 2003, however, also a potentially fatal, emerging human disease, severe acute respiratory syndrome (SARS), was shown to be caused by a coronavirus<sup>27,63</sup> (recently reviewed in reference<sup>76</sup>). Most nidoviruses have very restricted host specificity and infect only one or a narrow range of closely related species. One of the notable exceptions is SARS-CoV, which has been shown to infect a range of mammals<sup>45</sup>, a property that has been advantageous for the development of animal models for research on SARS-CoV pathogenesis<sup>81,99</sup>.

The different nidovirus groups encode widely different sets of structural proteins and, accordingly, have distinct virion architectures. Ronivirus and bafinivirus particles are rod-shaped and contain tubular helical nucleocapsids<sup>84,96</sup>. A helical nucleocapsid structure is also characteristic of the spherical coronavirus and, presumably, torovirus particle<sup>4,70,106</sup>, while arterivirus nucleocapsids are likely icosahedral (Fig. 1). The lipoprotein envelopes of nidovirus particles also differ significantly in morphology, largely due to the dissimilar arrays of viral envelope proteins they contain. The virions of *Coronaviridae* and *Roniviridae* representatives are characterized by large protrusions from the envelope surface formed by membrane-spanning viral proteins, while arterivirus particles carry relatively small envelope projections. The viral envelope proteins mediate attachment of nidovirus particles to receptors on the surface of host cells and fusion of the viral envelope with cellular membranes, an event that can take place either at the plasma or endosomal membranes. Following release of the viral nucleocapsid into the host cell cytoplasm, the nidovirus genome is uncoated and translated by host cell ribosomes. In the late stages of nidovirus infection, newly made genomic RNA molecules associate with nucleocapsid proteins. Interactions of the resulting nucleocapsid structures with intracellular membranes containing the viral envelope proteins trigger budding of progeny virions into the lumen of



**Figure 1.** Virion architecture in the order *Nidovirales*. Characteristic electron micrographs of negatively stained particles (left panels) and schematic diagrams (right panels) from representatives of the five main nidovirus genera: arterivirus, coronavirus, torovirus, bafinivirus and ronivirus. Scale bars represent 50 nm in the arterivirus image and 100 nm in all other images. Note the differences in nucleocapsid symmetry and envelope protein composition. Abbreviations: N, nucleocapsid protein; M, membrane protein; S, spike protein; E, envelope protein; GP or gp, glycoprotein, HE, hemagglutinin-esterase. GP<sub>5</sub> and M are the major glycoproteins in arterivirus particles, while GP<sub>2</sub>, GP<sub>3</sub>, and GP<sub>4</sub> are minor envelope components. Toroviruses, bafiniviruses, and roniviruses lack an equivalent of the E protein that is present in corona- and arteriviruses. Adapted from Enjuanes *et al.*<sup>32</sup> and Snijder *et al.*<sup>92</sup>

the smooth endoplasmic reticulum and/or the Golgi complex, followed by the release of progeny particles into extracellular space by exocytosis. The mechanisms of virus entry and assembly of infectious progeny employed by representatives of the major nidovirus groups are evidently different, which stands in contrast to similarities observed in other key viral properties. These include the order in which viral genes are encoded in the polycistronic +RNA genomes of nidoviruses, the mechanisms used to express those genes, and the composition of the nidovirus replicative machinery. Collectively, these observations argue for a common evolutionary ancestry of viruses united in the *Nidovirales* order.

## Arteriviridae



## Coronaviridae

### Coronavirinae



### Torovirinae



## Roniviridae



**Figure 2.** The polycistronic nature of nidovirus genomes. The genomic RNA of selected representatives of the *Arteriviridae*, *Coronaviridae* and *Roniviridae* families are drawn to the same scale. The 5'-proximal replicase open reading frames (ORFs), as well as the downstream ORFs encoding the viral structural proteins are depicted. The ORF1a/1b ribosomal frameshift site (RFS) and the 3' poly(A) tail ( $A_n$ ) are also indicated. ORFs encoding the viral replicase and viral structural proteins are designated (for abbreviations, see Fig. 1). Note the large size difference between the genomes of arteriviruses and other nidovirus families and subfamilies. EAV, equine arteritis virus; SARS-CoV, severe acute respiratory syndrome coronavirus; BToV, bovine torovirus; GAV, gill-associated virus.

The single-stranded RNA genomes of nidoviruses are 3' polyadenylated and carry a 5' cap structure. The genome size of *Arteriviridae*, ranging between 12.7 and 15.7 kb, differs considerably from those of *Coronaviridae* and *Roniviridae* - the virus groups with the largest and genetically most complex RNA genomes described to date (25-32 kb)<sup>39</sup>. The nonstructural proteins (nsps) of nidoviruses are encoded in a large "replicase" gene that occupies the 5'-proximal two-thirds to three-quarters of the genomic RNA and consists of two ORFs - ORF1a and ORF1b. A set of small ORFs, coding for the viral structural proteins and, in some nidovirus representatives, "accessory" proteins, are found in the 3'-proximal part of nidovirus genomes.

The nidovirus replicative cycle starts with translation of the genomic RNA, initiated at the ORF1a start codon following ribosomal scanning and resulting in the production of a large polyprotein, pp1a. A -1 ribosomal frameshift just upstream of the ORF1a termination codon results in the C-terminal extension of a fraction of pp1a with the ORF1b-encoded polypeptide, giving rise to pp1ab<sup>15</sup>. The two large replicase polyproteins are co- and post-translationally processed by virus-encoded (auto)proteinases into 13-16 individual nsps, which assemble into a cytoplasmic RNA-synthesizing complex that is anchored to modified cellular membranes. The structural protein ORFs, which are inaccessible to host cell ribosomes due to their 3'-proximal positions in the genomic RNA (Fig. 2), are expressed from a set of sg mRNAs. These molecules are 3'-coterminal with the viral genome, resulting in a so-called "nested" set of viral RNA species, a property reflected in the name of the

virus order (“nidus” is the Latin word for nest). The sg mRNAs of corona- and arteriviruses also contain a short sequence identical to the genomic 5' end<sup>5,22,64,95</sup>.

## THE NIDOVIRUS REPLICASE: AN ENZYME TREASURE CHEST

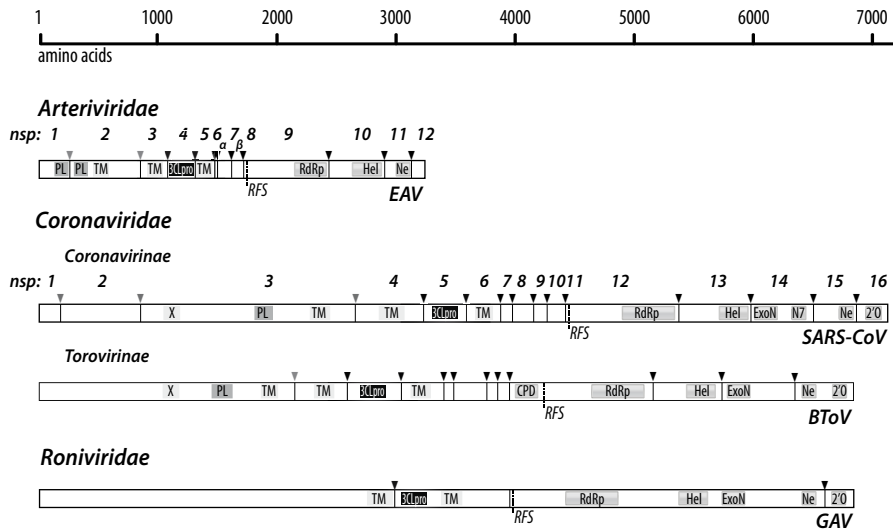
Despite the large size difference between the replicase pp1ab polyproteins, ranging from 3175 amino acids (aa) for the arterivirus equine arteritis virus (EAV) to ~7200 aa for the coronavirus murine hepatitis virus (MHV), the conserved array of replicative domains and their sequential arrangement in the large nidovirus replicase gene formed the basis for nidovirus unification (Fig. 3). The ORF1a-encoded replicase subunits include the “main” viral proteinase, an enzyme with a chymotrypsin-like fold and narrow substrate specificity. This enzyme mediates limited proteolysis of the C-terminal half of pp1a and the ORF1b-encoded portion of pp1ab at a number of sites<sup>1,6,24,42,66,94</sup>. Up to three ORF1a-encoded “accessory” proteinases are responsible for the (auto)catalytic processing of the N-terminal part of the nidovirus replicase polyproteins (reviewed in reference<sup>111</sup>). The replicase gene portion upstream of the ribosomal frameshift site also specifies a number of hydrophobic domains flanking the main proteinase. Transmembrane domain-containing nidovirus nsps have presumed or established roles in the rearrangement of host cell membranes into virus-induced compartments and the anchoring of nidovirus RNA-synthesizing complexes to these modified membranes in infected cells<sup>2,20,36,60,72-74,78,93</sup>.

### RdRp and helicase

ORF1b is the most conserved part of nidovirus genomes. The central viral enzymatic activities for RNA-templated RNA synthesis - the RdRp and helicase, are encoded in ORF1b, and phylogenetic analyses provided strong evidence for a common ancestry of these key enzymes in the different nidovirus groups<sup>21,24,41,42,48,91</sup>. The RdRp domain is found in the C-terminal portion of a larger replicase subunit whose size differs considerably among nidoviruses (Fig. 3). Recombinant forms of the RdRp-containing nsps from EAV (nsp9) and SARS-CoV (nsp12) have been recently produced, and their initial biochemical characterization revealed important functional differences. For example, EAV nsp9 is able to initiate RNA synthesis *de novo* but SARS-CoV nsp12 is a primer-dependent RdRp<sup>8,100</sup>. Although both proteins could catalyze RNA polymerization *in vitro* in the absence of other viral or cellular proteins, the recombinant EAV RdRp could not utilize sequences derived from the 3' end of the viral genome as a template, suggesting additional requirements for its activity *in vivo*. It is interesting to note that coronavirus nsp8 has been proposed to act as an “RNA primase”, based on its ability to synthesize short oligoribonucleotides *in vitro*<sup>49</sup>. No functional counterpart of this protein in the arterivirus replicase has been identified to date.

The availability of recombinant nidovirus RdRps provides a solid basis both for the detailed functional characterization of their enzymatic activities that should yield valuable





**Figure 3.** Organization and composition of nidovirus replicases. Currently mapped functional domains in the pp1ab replicase polyprotein of the *Arteriviridae*, *Coronaviridae* and *Roniviridae* representatives listed in Fig. 2 are shown schematically to scale. The border between amino acids encoded in ORF1a and ORF1b is indicated as RFS (ribosomal frameshift). Arrowheads represent cleavage sites processed by virus-encoded proteinases. Sites cleaved by papain-like (PL) accessory proteinases are shown in grey, while those processed by the 3C-like main proteinase (3CLpro) are depicted in black. The resulting nonstructural proteins (nsp) are numbered where cleavage site maps of pp1ab proteins are available. In addition to the proteinase domains, the location of putative transmembrane domains (TM) and the major viral enzymatic domains are shown: RNA-dependent RNA polymerase (RdRp), helicase (Hel), exoribonuclease (ExoN), endoribonuclease (Ne, NendoU in the main text), N7-methyltransferase (N7), 2'-O-methyltransferase (2'O), ADP-ribose-1"-phosphatase (X) and cyclic phosphodiesterase (CPD).

mechanistic insights into nidovirus RNA synthesis. Of particular interest would be the significance of the nidovirus helicase for viral RNA synthesis. RNA helicases are a diverse class of enzymes that unwind RNA duplexes using the energy of ATP hydrolysis. In nidoviruses, the helicase domain, much like the RdRp domain, is part of a larger replicase subunit (Fig. 3), which also contains an N-terminal predicted zinc-binding domain with 13 conserved Cys and His residues<sup>42,104</sup>. Recombinant forms of the arteri- and coronavirus nsps comprising the helicase domain (nsp10 and nsp13, respectively) have ATPase activities and can unwind RNA and DNA duplexes in a 5'-to-3' direction *in vitro*<sup>51,52,86,88</sup>. The helicase activity of coronavirus nsp13 is remarkably processive, permitting strand separation of long stretches of double-stranded nucleic acids, and the protein also exhibits RNA 5'-triphosphatase activity *in vitro*<sup>51,52</sup>. The zinc-binding domains of arterivirus nsp10 and coronavirus nsp13 are also critical for the *in vitro* ATPase and helicase activities of the proteins<sup>87</sup>. Although residues from this domain are unlikely to be involved in catalysis, zinc coordination might assist the proper folding of the entire replicase subunit and/or mediate interactions of the protein with substrate RNA molecules. Notably, the 5'-to-3'

polarity of nidovirus helicases has not been reconciled so far with their presumed role in unwinding local double-stranded RNA structures that might hinder the progress of the RdRp during viral RNA synthesis, which proceeds in the opposite direction.

### **The nidovirus endoribonuclease (NendoU)**

In addition to the core viral enzymes, the replicative machinery of nidoviruses includes several subunits with rather unusual RNA-processing activities that have few or no known counterparts in the RNA virus world. The majority of these enzymes are encoded only by nidoviruses with large genomes, but one of them - the nidovirus endoribonuclease (NendoU) domain, is conserved throughout the nidovirus order<sup>39,89</sup>. To date, no NendoU counterparts have been identified in other RNA viruses, and this domain is thus considered a genetic marker of *Nidovirales*. The NendoU domain is encoded in ORF1b, downstream of the RdRp and helicase, and is N-terminally fused to another domain in a larger replicase subunit - nsp11 in arteriviruses and nsp15 in coronaviruses (Fig. 3). Its endoribonuclease function was originally predicted based on distant sequence homology between the NendoU domain and a small family of prokaryotic and eukaryotic proteins prototyped by XendoU, an endoribonuclease from *Xenopus laevis*<sup>89</sup>. Several coronavirus nsp15 orthologs were shown to cleave RNA 3' of uridylates *in vitro*, and the enzyme's critical role in the coronavirus replicative cycle was established using recombinant coronaviruses that expressed mutant forms of nsp15<sup>50,59</sup>. Biochemical and structural studies have provided some insights into the catalytic activity of coronavirus NendoU domains and the determinants of the enzyme's substrate specificity<sup>10-12,46,50,55,80</sup>. The cognate substrate of NendoU in infected cells has not been identified, however, and the molecular details of the role of NendoU in the nidovirus replicative cycle, as well as the mechanism(s) to protect viral (and possibly cellular) RNA molecules in infected cells from rapid degradation, remain unexplored thus far.

### **Exoribonuclease, 2' O-methyltransferase and N7- methyltransferase**

Apart from NendoU, members of the *Coronaviridae* and *Roniviridae* encode two additional conserved enzymatic domains in the 3'-proximal region of ORF1b – a 3'-to-5' exoribonuclease (ExoN) and a (nucleoside-2' O)-methyltransferase (2' O-MTase), mapping to coronavirus nsp14 and nsp16, respectively<sup>23,39,68,89</sup>. A role of ExoN in improving the low fidelity of RdRp-mediated RNA synthesis, and thus contributing to the maintenance of the large genomes in these two nidovirus families, was postulated<sup>39,89</sup>. ExoN inactivation by substitution of active-site residues correlates with severe defects in coronavirus RNA synthesis<sup>28,68</sup>. MHV ExoN mutants were shown to accumulate 15-fold more mutations than the wild-type virus<sup>28</sup>, but rigorous experimental proof of an ExoN role in proofreading is currently lacking. Interestingly, a (guanine-N7)-methyltransferase (N7-MTase) activity of nsp14 was recently identified in a yeast genetic screen for cap-forming enzymes of coronaviruses. The ExoN and N7-MTase activities of nsp14 seem to be functionally separated

in the protein's primary structure<sup>18</sup>. Nidoviruses with large genomes thus possess at least two enzymes that could participate in the formation of 5'-terminal cap 1 structures. Recently, mRNA cap methylation was reconstituted *in vitro* using recombinant nsp14 and nsp16 from SARS-CoV. This study also reported a surprising role of nsp10 as a critical factor for the 2'O-MTase activity of nsp16 *in vitro*<sup>14</sup>. The coronavirus helicase was also suggested to play a role in viral mRNA capping due to its RNA 5'-triphosphatase activity, since removal of the 5'  $\gamma$ -phosphate is the first step in cap formation<sup>52</sup>. By contrast, no RNA capping functions have been identified in the arterivirus replicase to date.

### ADP-ribose-1''-phosphatase and cyclic phosphodiesterase

The repertoire of known and/or predicted nidovirus RNA-processing enzymes is completed by two domains - a cyclic phosphodiesterase (CPD) and an ADP-ribose-1''-phosphatase (ADRP), which are encoded by a smaller subset of nidovirus groups in genomic regions other than ORF1b<sup>89</sup>. Both domains are not unique to nidovirus representatives - an ADRP domain is found in the replicase polyproteins of all mammalian viruses of the alphavirus-like supergroup<sup>43</sup>, while a distantly related CPD domain can be found in the genomes of some dsRNA viruses<sup>89</sup>. The CPD domain-coding sequence is found in the 3' end of ORF1a just upstream of the ribosomal frameshift in *Torovirus* representatives, while in a subset of *Betacoronavirus* members it is not encoded in the replicase gene, but expressed from a sg mRNA and called the ns2 protein<sup>90</sup> (see Fig. 3). None of the other currently known nidoviruses encodes a CPD homolog. Cyclic phosphodiesterase activity of this domain has not been demonstrated so far, and deletion of the ns2-coding sequence from the MHV genome does not affect virus replication in cell culture<sup>85</sup>. However, a number of amino acid substitutions in ns2, including replacements of predicted catalytic CPD residues, were found to attenuate virus growth in mice, particularly in the livers of infected animals<sup>82,97</sup>.

A similar phenotype has been reported for an MHV mutant with a substitution in the ADRP active site, which replicated efficiently both in cultured cells and in the livers of infected mice, but did not induce liver disease<sup>33</sup>. The ADRP domain was originally identified as X domain<sup>89</sup> and, subsequently, macro domain, as it is evolutionarily related to a conserved domain found in the C-terminal region of the histone macroH2A<sup>75</sup>. All representatives of the *Coronaviridae* family harbor a macro domain in the N-terminal region of pp1a/pp1ab, and in coronaviruses this domain is part of the largest replicase subunit - nsp3. The biochemical characterization of coronavirus macro domains was initially focused on its ability to convert ADP-ribose-1''-phosphate, a by-product of cellular pre-tRNA splicing, to ADP-ribose<sup>79</sup>. Recombinant coronavirus macro domains exhibited specific, but poor ADRP activity *in vitro*, and recent reports have suggested these domains bind efficiently to poly(ADP-ribose) - a unique post-translational protein modification involved in multiple cell signaling pathways, among which those related to cell survival<sup>83</sup>. It is unclear, however, whether ADP-ribose binding is a conserved property of coronavirus macro domains, since a recombinant form of this domain from the group

3 coronavirus infectious bronchitis virus (IBV) failed to bind ADP-ribose *in vitro*<sup>77</sup>. Curiously, this was a property only of the macro domain derived from IBV strain Beaudette, while the IBV strain M41 macro domain has been successfully crystallized in complex with ADP-ribose<sup>107</sup>. Taken together, the properties of the CPD and ADRP domains are most consistent with roles of these enzymatic activities in the modulation of host cellular responses to viral infection<sup>30,33,79,82</sup>.

### **ORF1a-encoded replicase subunits**

In contrast to the clear evolutionary relationship among ORF1b-encoded domains, the N-proximal parts of pp1a/pp1ab are very poorly conserved among nidovirus families<sup>21,24,90,110</sup>. Aside from the identical sequential arrangement of membrane-associated nsps and the main viral proteinase, another discernible common feature is the presence of one to three “accessory” proteinases that direct the autocatalytic processing of the N-proximal replicase polyprotein regions in which they reside (reviewed in <sup>111</sup>). The papain-like cysteine proteinase domains residing in nsp2 of arteriviruses and nsp3 of coronaviruses reportedly can hydrolyze ubiquitin and interferon-stimulated gene product 15 (ISG15) from covalently modified proteins. These activities are due to the narrow substrate specificities of these proteinases, which match closely those of known cellular deubiquitinases. Consequently, papain-like proteinases of arteri- and coronaviruses have been proposed to play a role in counteracting ubiquitin- and ISG15-dependent cellular antiviral responses<sup>7,35,65</sup>. Elegant biochemical studies on SARS-CoV nsp1, a protein with no known enzymatic activities, have suggested this replicase subunit could also be involved in suppressing antiviral signaling pathways by inhibiting host translation and promoting host mRNA degradation<sup>57,58,69</sup>. This function of nsp1, however, is dispensable for SARS-CoV replication in cultured cells<sup>69</sup> and might only be important for pathogenicity in the natural host, as demonstrated for its ortholog in MHV<sup>112</sup>.

In contrast to the apparent expendability of coronavirus nsp1 and nsp2 for viral RNA replication in cell culture<sup>16,25,44</sup>, the similarly positioned region of the arterivirus replicase polyprotein contains a protein factor – nsp1, which is critical for viral RNA synthesis. EAV nsp1 is not essential for genome replication, but absolutely required for the synthesis of sg mRNAs, and a similar function has been proposed for nsp1 $\alpha$  from PRRSV<sup>61,101</sup>. Interestingly, PRRSV nsp1 $\alpha$  and nsp1 $\beta$  have recently been reported to function as interferon antagonists<sup>9,19</sup>. Multiple ORF1a-encoded nsps from different nidoviruses thus may contribute to the evasion of innate immunity surveillance mechanisms, a function they might be exceptionally well suited for, considering they are the first viral proteins produced upon infection.

The ORF1a gene products thus seem to play diverse roles in viral RNA synthesis, providing subunits that modify the intracellular environment for this process to take place and ensure the regulated proteolysis of replicase polyproteins. It is clear that ORF1a replicase gene products can also have key regulatory functions in viral RNA synthesis, as demon-

strated for arterivirus nsp1<sup>61,101</sup> and proposed for the nsp7 to nsp10 replicase subunits of coronaviruses<sup>26,29,67,98,109,113</sup>.

## SCOPE AND OUTLINE OF THIS THESIS

The functional analysis of nidovirus replicase subunits, both *in vitro* and in the context of viral infection, is imperative for addressing fundamental questions about the composition of nidovirus RNA-synthesizing complexes and the regulation of their different activities. This information is also of significance for the design of selective inhibitors of nidovirus replication. The availability of full-length cDNA clones greatly facilitates the integration of results derived from biochemical studies on nidovirus replicase subunits and their functional analyses in the context of viral replication. The development of a reverse genetics system for EAV, which has the smallest nidovirus genome and also replicates robustly in a variety of cell lines without the need for extreme safety measures, has made this virus exceptionally suited for functional dissection of the nidovirus replicase. Part of the work described in this thesis involved the characterization of the conserved NendoU domain, mapping to arterivirus nsp11, and its significance for the EAV replicative cycle. Another part of this thesis focused on nsp1, previously identified as a candidate “molecular switch” between genome replication and sg mRNA synthesis in EAV, and was aimed at gaining more insight into the regulatory role of the protein in the poorly-understood process of discontinuous minus-strand synthesis.

Chapter 2 explores the inherent multifunctionality of +RNA virus genomes. Well-documented control mechanisms that determine the use of these molecules as templates for translation, minus-strand synthesis or encapsidation, as well as the high degree of integration between these processes in different +RNA virus families are discussed. The current knowledge on such regulatory events in the replicative cycle of nidoviruses is summarized.

Chapter 3 establishes the biological importance of the conserved NendoU domain by a site-directed mutagenesis approach. Deletion of the NendoU domain from EAV nsp11 and certain amino acid replacements rendered viral RNA accumulation undetectable when engineered in the full-length EAV cDNA clone. Substitutions of proposed catalytic residues, however, resulted in viable mutants with greatly reduced infectious progeny titers and small-plaque phenotypes. A more detailed analysis of the latter mutants revealed defects in viral RNA accumulation and identified a potential link between the NendoU domain and viral sg mRNA synthesis.

Chapter 4 describes the biochemical characterization of nsp11 from two distantly related arteriviruses, EAV and PRRSV. Bacterially expressed nsp11 proteins were found to efficiently hydrolyze RNA 3' of pyrimidines *in vitro*, with a modest preference for cleavage at uridylates. Comparative analysis of arterivirus nsp11 and coronavirus nsp15 purified under identical conditions revealed common and distinct features of the endoribonucleic activities of these nidovirus replicase subunits that carry distantly related NendoU

domains. The biological significance of this remarkable RNA-processing activity, likely common to all nidoviruses, is discussed.

Chapter 5 addresses the significance of nsp1's autoproteolytic release from the replicase polyproteins and the importance of the protein's ZF domain in the context of the complete replicative cycle of EAV. Cleavage of the nsp1/2 site was found to be a prerequisite for EAV RNA synthesis. By contrast, several substitutions of predicted zinc-coordinating residues selectively blocked sg mRNA accumulation, while genome replication was either not affected or even increased. Other mutations of predicted zinc-binding residues, however, did not affect viral RNA synthesis considerably but severely reduced the production of infectious progeny particles. Apart from playing key roles in replicase maturation and sg mRNA synthesis, nsp1 was thus also found to have a critical function in virion biogenesis.

Chapter 6 presents a study in which both reverse and forward genetics were employed to gain more insight into the multiple regulatory roles of EAV nsp1. Alanine replacement of non-conserved clusters of polar residues found throughout the nsp1 sequence was used as an approach aimed at expanding the repertoire of viable EAV nsp1 mutants with discernible defects in one or more of the protein's functions. The analysis of mutant and pseudorevertant phenotypes revealed that the relative abundance of EAV mRNAs is tightly controlled by an intricate network of interactions involving all nsp1 subdomains. Moreover, nsp1 was implicated in modulating the accumulation of full-length and subgenome-length minus-strand templates for viral mRNA synthesis. The quantitative balance among viral mRNA species was also found to be critical for efficient production of new virus particles. These data establish nsp1 as a major coordinator of the EAV replicative cycle.

Chapter 7 describes a protocol for the expression and purification of recombinant EAV nsp1 from *E. coli*. The initial difficulties encountered in attempting to obtain recombinant nsp1 were found to be largely related to the protein's poor solubility, and a range of fusion tags and expression conditions were explored in order to overcome this problem. Expression of the protein from a construct that allowed its autoproteolytic release from a precursor polypeptide in *E. coli* was found to be critical for the successful purification of soluble, stable recombinant nsp1.

Chapter 8 discusses the main results described in this thesis and their implications. Potential approaches for future research are also outlined.

## REFERENCES

1. **Anand, K., G. J. Palm, J. R. Mesters, S. G. Siddell, J. Ziebuhr, and R. Hilgenfeld.** 2002. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *EMBO J.* **21**:3213-3224
2. **Baliji, S., S. A. Cammer, B. Sobral, and S. C. Baker.** 2009. Detection of nonstructural protein 6 in murine coronavirus-infected cells and analysis of the transmembrane topology by using bioinformatics and molecular approaches. *J.Virol.* **83**:6957-6962
3. **Ban, N., P. Nissen, J. Hansen, P. B. Moore, and T. A. Steitz.** 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**:905-920
4. **Barcena, M., G. T. Oostergetel, W. Bartelink, F. G. Faas, A. Verkleij, P. J. Rottier, A. J. Koster, and B. J. Bosch.** 2009. Cryo-electron tomography of mouse hepatitis virus: Insights into the structure of the coronavirus. *Proc.Natl.Acad.Sci.U.S.A* **106**:582-587
5. **Baric, R. S., S. A. Stohman, and M. M. Lai.** 1983. Characterization of replicative intermediate RNA of mouse hepatitis virus: presence of leader RNA sequences on nascent chains. *J.Virol.* **48**:633-640
6. **Barrette-Ng, I. H., K. K. S. Ng, B. L. Mark, D. van Aken, M. M. Cherney, C. Garen, Y. Kolodenco, A. E. Gorbalenya, E. J. Snijder, and M. N. G. James.** 2002. Structure of arterivirus nsp4 - The smallest chymotrypsin-like proteinase with an alpha/beta C-terminal extension and alternate conformations of the oxyanion hole. *Journal of Biological Chemistry* **277**:39960-39966
7. **Barretto, N., D. Jukneliene, K. Ratia, Z. Chen, A. D. Mesecar, and S. C. Baker.** 2005. The papain-like protease of severe acute respiratory syndrome coronavirus has deubiquitinating activity. *J.Virol.* **79**:15189-15198
8. **Beerens, N., B. Selisko, S. Ricagno, I. Imbert, L. van der Zanden, E. J. Snijder, and B. Canard.** 2007. De novo initiation of RNA synthesis by the arterivirus RNA-dependent RNA polymerase. *J.Virol.* **81**:8384-8395
9. **Beura, L. K., S. N. Sarkar, B. Kwon, S. Subramaniam, C. Jones, A. K. Pattnaik, and F. A. Osorio.** 2010. Porcine reproductive and respiratory syndrome virus nonstructural protein 1beta modulates host innate immune response by antagonizing IRF3 activation. *J.Virol.* **84**:1574-1584
10. **Bhardwaj, K., L. Guarino, and C. C. Kao.** 2004. The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. *J.Virol.* **78**:12218-12224
11. **Bhardwaj, K., S. Palaninathan, J. M. Alcantara, L. L. Yi, L. Guarino, J. C. Sacchettini, and C. C. Kao.** 2008. Structural and functional analyses of the severe acute respiratory syndrome coronavirus endoribonuclease Nsp15. *J.Biol.Chem.* **283**:3655-3664
12. **Bhardwaj, K., J. Sun, A. Holzenburg, L. A. Guarino, and C. C. Kao.** 2006. RNA recognition and cleavage by the SARS coronavirus endoribonuclease. *J.Mol.Biol.* **361**:243-256
13. **Bieling, P., M. Beringer, S. Adio, and M. V. Rodnina.** 2006. Peptide bond formation does not involve acid-base catalysis by ribosomal residues. *Nat.Struct.Mol.Biol.* **13**:423-428
14. **Bouvet, M., C. Debarnot, I. Imbert, B. Selisko, E. J. Snijder, B. Canard, and E. Decroly.** 2010. In vitro reconstitution of SARS-Coronavirus mRNA cap methylation. *PLoS.Pathog.*, in press
15. **Brierley, I., P. Digard, and S. C. Inglis.** 1989. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* **57**:537-547
16. **Brockway, S. M. and M. R. Denison.** 2005. Mutagenesis of the murine hepatitis virus nsp1-coding region identifies residues important for protein processing, viral RNA synthesis, and viral replication. *Virology* **340**:209-223
17. **Cavanagh, D.** 1997. Nidovirales: a new order comprising Coronaviridae and Arteriviridae. *Arch. Virol.* **142**:629-633

18. **Chen, Y., H. Cai, J. Pan, N. Xiang, P. Tien, T. Ahola, and D. Guo.** 2009. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc.Natl. Acad.Sci.U.S.A* **106**:3484-3489
19. **Chen, Z., S. Lawson, Z. Sun, X. Zhou, X. Guan, J. Christopher-Hennings, E. A. Nelson, and Y. Fang.** 2010. 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* **398**:87-97
20. **Clementz, M. A., A. Kanjanahaluethai, T. E. O'Brien, and S. C. Baker.** 2008. Mutation in murine coronavirus replication protein nsp4 alters assembly of double membrane vesicles. *Virology* **375**:118-129
21. **Cowley, J. A., C. M. Dimmock, K. M. Spann, and P. J. Walker.** 2000. Gill-associated virus of *Penaeus monodon* prawns: an invertebrate virus with ORF1a and ORF1b genes related to arteri- and coronaviruses. *J.Gen.Virol.* **81**:1473-1484
22. **de Vries, A. A., E. D. Chirnside, P. J. Bredenbeek, L. A. Gravestein, M. C. Horzinek, and W. J. Spaan.** 1990. All subgenomic mRNAs of equine arteritis virus contain a common leader sequence. *Nucleic Acids Res.* **18**:3241-3247
23. **Decroly, E., I. Imbert, B. Coutard, M. Bouvet, B. Selisko, K. Alvarez, A. E. Gorbalenya, E. J. Snijder, and B. Canard.** 2008. Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2'O)-methyltransferase activity. *J.Virol.* **82**:8071-8084
24. **den Boon, J. A., E. J. Snijder, E. D. Chirnside, A. A. de Vries, M. C. Horzinek, and W. J. Spaan.** 1991. Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. *J.Virol.* **65**:2910-2920
25. **Denison, M. R., B. Yount, S. M. Brockway, R. L. Graham, A. C. Sims, X. Lu, and R. S. Baric.** 2004. Cleavage between replicase proteins p28 and p65 of mouse hepatitis virus is not required for virus replication. *J.Virol.* **78**:5957-5965
26. **Donaldson, E. F., A. C. Sims, R. L. Graham, M. R. Denison, and R. S. Baric.** 2007. Murine hepatitis virus replicase protein nsp10 is a critical regulator of viral RNA synthesis. *J.Virol.* **81**:6356-6368
27. **Drosten, C., S. Gunther, W. Preiser, S. van der Werf, H. R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R. A. Fouchier, A. Berger, A. M. Burguiere, J. Cinatl, M. Eickmann, N. Escriou, K. Grywna, S. Kramme, J. C. Manuguerra, S. Muller, V. Rickerts, M. Sturmer, S. Vieth, H. D. Klenk, A. D. Osterhaus, H. Schmitz, and H. W. Doerr.** 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N.Engl.J.Med.* **348**:1967-1976
28. **Eckerle, L. D., X. Lu, S. M. Sperry, L. Choi, and M. R. Denison.** 2007. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J.Virol.* **81**:12135-12144
29. **Egloff, M. P., F. Ferron, V. Campanacci, S. Longhi, C. Rancurel, H. Dutartre, E. J. Snijder, A. E. Gorbalenya, C. Cambillau, and B. Canard.** 2004. The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. *Proc.Natl.Acad.Sci.U.S.A* **101**:3792-3796
30. **Egloff, M. P., H. Malet, A. Putics, M. Heinonen, H. Dutartre, A. Frangeul, A. Gruez, V. Campanacci, C. Cambillau, J. Ziebuhr, T. Ahola, and B. Canard.** 2006. Structural and functional basis for ADP-ribose and poly(ADP-ribose) binding by viral macro domains. *J.Virol.* **80**:8493-8502
31. **Eigen, M.** 1993. The origin of genetic information: viruses as models. *Gene* **135**:37-47
32. **Enjuanes, L., A. E. Gorbalenya, R. J. de Groot, J. A. Cowley, J. Ziebuhr, and E. J. Snijder.** 2008. *Nidovirales*, p. 419-430. In: B. W. J. Mahy and M. H. V. an Regenmortel (eds.), *Encyclopedia of Virology*, 5 vols. Elsevier, Oxford.



33. **Eriksson, K. K., L. Cervantes-Barragan, B. Ludewig, and V. Thiel.** 2008. Mouse hepatitis virus liver pathology is dependent on ADP-ribose-1<sup>st</sup>-phosphatase, a viral function conserved in the alpha-like supergroup. *J.Virol.* **82**:12325-12334
34. **Flegel, T. W.** 1997. Major viral diseases of the black tiger prawn (*Penaeus monodon*) in Thailand. *World J. Microbiol. Biotechnol.* **13**:433-442
35. **Frias-Staheli, N., N. V. Giannakopoulos, M. Kikkert, S. L. Taylor, A. Bridgen, J. Paragas, J. A. Richt, R. R. Rowland, C. S. Schmaljohn, D. J. Lenschow, E. J. Snijder, A. Garcia-Sastre, and H. W. Virgin.** 2007. Ovarian tumor domain-containing viral proteases evade ubiquitin- and ISG15-dependent innate immune responses. *Cell Host & Microbe* **2**:404-416
36. **Gadlage, M. J., J. S. Sparks, D. C. Beachboard, R. G. Cox, J. D. Doyle, C. C. Stobart, and M. R. Denison.** 2010. Murine hepatitis virus nonstructural protein 4 regulates virus-induced membrane modifications and replication complex function. *J.Virol.* **84**:280-290
37. **Gilbert, W.** 1986. Origin of life: The RNA world. *Nature* **319**:618
38. **Gonzalez, J. M., P. Gomez-Puertas, D. Cavanagh, A. E. Gorbalenya, and L. Enjuanes.** 2003. A comparative sequence analysis to revise the current taxonomy of the family Coronaviridae. *Arch. Virol.* **148**:2207-2235
39. **Gorbalenya, A. E., L. Enjuanes, J. Ziebuhr, and E. J. Snijder.** 2006. Nidovirales: evolving the largest RNA virus genome. *Virus Res.* **117**:17-37
40. **Gorbalenya, A. E. and E. V. Koonin.** 1989. Viral proteins containing the purine NTP-binding sequence pattern. *Nucleic Acids Res.* **17**:8413-8440
41. **Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov.** 1988. A novel superfamily of nucleoside triphosphate-binding motif containing proteins which are probably involved in duplex unwinding in DNA and RNA replication and recombination. *FEBS Lett.* **235**:16-24
42. **Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov.** 1989. Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Res.* **17**:4847-4861
43. **Gorbalenya, A. E., E. V. Koonin, and M. M. Lai.** 1991. Putative papain-related thiol proteases of positive-strand RNA viruses. Identification of rubi- and aphthovirus proteases and delineation of a novel conserved domain associated with proteases of rubi-, alpha- and coronaviruses. *FEBS Lett.* **288**:201-205
44. **Graham, R. L., A. C. Sims, S. M. Brockway, R. S. Baric, and M. R. Denison.** 2005. The nsp2 replicase proteins of murine hepatitis virus and severe acute respiratory syndrome coronavirus are dispensable for viral replication. *J.Virol.* **79**:13399-13411
45. **Groneberg, D. A., S. M. Poutanen, D. E. Low, H. Lode, T. Welte, and P. Zabel.** 2005. Treatment and vaccines for severe acute respiratory syndrome. *Lancet Infect.Dis.* **5**:147-155
46. **Guarino, L. A., K. Bhardwaj, W. Dong, J. Sun, A. Holzenburg, and C. Kao.** 2005. Mutational analysis of the SARS virus Nsp15 endoribonuclease: identification of residues affecting hexamer formation. *J.Mol.Biol.* **353**:1106-1117
47. **Guerrier-Takada, C., K. Gardiner, T. Marsh, N. Pace, and S. Altman.** 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**:849-857
48. **Hodgman, T. C.** 1988. A new superfamily of replicative proteins. *Nature* **333**:22-23
49. **Imbert, I., J. C. Guillemot, J. M. Bourhis, C. Bussetta, B. Coutard, M. P. Egloff, F. Ferron, A. E. Gorbalenya, and B. Canard.** 2006. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J.* **25**:4933-4942
50. **Ivanov, K. A., T. Hertzog, M. Rozanov, S. Bayer, V. Thiel, A. E. Gorbalenya, and J. Ziebuhr.** 2004. Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc.Natl.Acad. Sci.U.S.A* **101**:12694-12699

51. **Ivanov, K. A., V. Thiel, J. C. Dobbe, Y. van der Meer, E. J. Snijder, and J. Ziebuhr.** 2004. Multiple enzymatic activities associated with Severe acute respiratory syndrome coronavirus helicase. *J.Virol.* **78**:5619-5632
52. **Ivanov, K. A. and J. Ziebuhr.** 2004. Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5'-triphosphatase activities. *J.Virol.* **78**:7833-7838
53. **Jeffares, D. C., A. M. Poole, and D. Penny.** 1998. Relics from the RNA world. *J.Mol.Evol.* **46**:18-36
54. **Johnston, W. K., P. J. Unrau, M. S. Lawrence, M. E. Glasner, and D. P. Bartel.** 2001. RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science* **292**:1319-1325
55. **Joseph, J. S., K. S. Saikatendu, V. Subramanian, B. W. Neuman, M. J. Buchmeier, R. C. Stevens, and P. Kuhn.** 2007. Crystal structure of a monomeric form of severe acute respiratory syndrome coronavirus endonuclease nsp15 suggests a role for hexamerization as an allosteric switch. *J.Virol.* **81**:6700-6708
56. **Joyce, G. F.** 1989. RNA evolution and the origins of life. *Nature* **338**:217-224
57. **Kamitani, W., C. Huang, K. Narayanan, K. G. Lokugamage, and S. Makino.** 2009. A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. *Nat.Struct.Mol.Biol.* **16**:1134-1140
58. **Kamitani, W., K. Narayanan, C. Huang, K. Lokugamage, T. Ikegami, N. Ito, H. Kubo, and S. Makino.** 2006. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. *Proc.Natl.Acad.Sci.U.S.A* **103**:12885-12890
59. **Kang, H., K. Bhardwaj, Y. Li, S. Palaninathan, J. Sacchettini, L. Guarino, J. L. Leibowitz, and C. C. Kao.** 2007. Biochemical and genetic analyses of murine hepatitis virus Nsp15 endoribonuclease. *J.Virol.* **81**:13587-13597
60. **Knoops, K., M. Kikkert, S. H. Worm, J. C. Zevenhoven-Dobbe, Y. van der Meer, A. J. Koster, A. M. Mommaas, and E. J. Snijder.** 2008. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS.Biol.* **6**:e226
61. **Kroese, M. V., J. C. Zevenhoven-Dobbe, J. N. Bos-de Ruijter, B. P. Peeters, J. J. Meulenberg, L. A. Cornelissen, and E. J. Snijder.** 2008. The nsp1alpha and nsp1 papain-like autoproteases are essential for porcine reproductive and respiratory syndrome virus RNA synthesis. *J.Gen.Virol.* **89**:494-499
62. **Kruger, K., P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, and T. R. Cech.** 1982. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **31**:147-157
63. **Ksiazek, T. G., D. Erdman, C. S. Goldsmith, S. R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J. A. Comer, W. Lim, P. E. Rollin, S. F. Dowell, A. E. Ling, C. D. Humphrey, W. J. Shieh, J. Guarner, C. D. Paddock, P. Rota, B. Fields, J. DeRisi, J. Y. Yang, N. Cox, J. M. Hughes, J. W. LeDuc, W. J. Bellini, and L. J. Anderson.** 2003. A novel coronavirus associated with severe acute respiratory syndrome. *N.Engl.J.Med.* **348**:1953-1966
64. **Lai, M. M., R. S. Baric, P. R. Brayton, and S. A. Stohlman.** 1984. Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus. *Proc.Natl. Acad.Sci.U.S.A* **81**:3626-3630
65. **Lindner, H. A., N. Fotouhi-Ardakani, V. Lytvyn, P. Lachance, T. Sulea, and R. Menard.** 2005. The papain-like protease from the severe acute respiratory syndrome coronavirus is a deubiquitinating enzyme. *J.Virol.* **79**:15199-15208

66. **Liu, D. X., I. Brierley, K. W. Tibbles, and T. D. Brown.** 1994. A 100-kilodalton polypeptide encoded by open reading frame (ORF) 1b of the coronavirus infectious bronchitis virus is processed by ORF 1a products. *J.Virol.* **68**:5772-5780
67. **Miknis, Z. J., E. F. Donaldson, T. C. Umland, R. A. Rimmer, R. S. Baric, and L. W. Schultz.** 2009. Severe acute respiratory syndrome coronavirus nsp9 dimerization is essential for efficient viral growth. *J.Virol.* **83**:3007-3018
68. **Minskaia, E., T. Hertzog, A. E. Gorbalenya, V. Campanacci, C. Cambillau, B. Canard, and J. Ziebuhr.** 2006. Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc.Natl.Acad.Sci.U.S.A* **103**:5108-5113
69. **Narayanan, K., C. Huang, K. Lokugamage, W. Kamitani, T. Ikegami, C. T. Tseng, and S. Makino.** 2008. Severe acute respiratory syndrome coronavirus nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. *J.Virol.* **82**:4471-4479
70. **Neuman, B. W., B. D. Adair, M. Yeager, and M. J. Buchmeier.** 2008. Purification and electron cryomicroscopy of coronavirus particles. *Methods Mol.Biol.* **454**:129-136
71. **Neumann, E. J., J. B. Kliebenstein, C. D. Johnson, J. W. Mabry, E. J. Bush, A. H. Seitzinger, A. L. Green, and J. J. Zimmerman.** 2005. Assessment of the economic impact of porcine reproductive and respiratory syndrome on swine production in the United States. *J.Am.Vet.Med.Assoc.* **227**:385-392
72. **Oostra, M., M. C. Hagemeijer, G. M. van, C. P. Bekker, E. G. te Lintelo, P. J. Rottier, and C. A. de Haan.** 2008. Topology and membrane anchoring of the coronavirus replication complex: not all hydrophobic domains of nsp3 and nsp6 are membrane spanning. *J.Virol.* **82**:12392-12405
73. **Oostra, M., E. G. te Lintelo, M. Deijs, M. H. Verheije, P. J. Rottier, and C. A. de Haan.** 2007. Localization and membrane topology of coronavirus nonstructural protein 4: involvement of the early secretory pathway in replication. *J.Virol.* **81**:12323-12336
74. **Pedersen, K. W., Y. van der Meer, N. Roos, and E. J. Snijder.** 1999. 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.* **73**:2016-2026
75. **Pehrson, J. R. and R. N. Fuji.** 1998. Evolutionary conservation of histone macroH2A subtypes and domains. *Nucleic Acids Res.* **26**:2837-2842
76. **Perlman, S. and J. Netland.** 2009. Coronaviruses post-SARS: update on replication and pathogenesis. *Nat.Rev.Microbiol.* **7**:439-450
77. **Piotrowski, Y., G. Hansen, Boomaars-van der Zanden AL, E. J. Snijder, A. E. Gorbalenya, and R. Hilgenfeld.** 2009. Crystal structures of the X-domains of a Group-1 and a Group-3 coronavirus reveal that ADP-ribose-binding may not be a conserved property. *Protein Sci.* **18**:6-16
78. **Posthuma, C. C., K. W. Pedersen, Z. Lu, R. G. Joosten, N. Roos, J. C. Zevenhoven-Dobbe, and E. J. Snijder.** 2008. Formation of the arterivirus replication/transcription complex: a key role for nonstructural protein 3 in the remodeling of intracellular membranes. *J.Virol.* **82**:4480-4491
79. **Putics, A., W. Filipowicz, J. Hall, A. E. Gorbalenya, and J. Ziebuhr.** 2005. ADP-ribose-1"-mono-phosphatase: a conserved coronavirus enzyme that is dispensable for viral replication in tissue culture. *J.Virol.* **79**:12721-12731
80. **Ricagno, S., M. P. Egloff, R. Ulferts, B. Coutard, D. Nurizzo, V. Campanacci, C. Cambillau, J. Ziebuhr, and B. Canard.** 2006. Crystal structure and mechanistic determinants of SARS coronavirus nonstructural protein 15 define an endoribonuclease family. *Proc.Natl.Acad.Sci.U.S.A* **103**:11892-11897
81. **Roberts, A., D. Deming, C. D. Paddock, A. Cheng, B. Yount, L. Vogel, B. D. Herman, T. Sheahan, M. Heise, G. L. Genrich, S. R. Zaki, R. Baric, and K. Subbarao.** 2007. A mouse-adapted SARS-coronavirus causes disease and mortality in BALB/c mice. *PLoS.Pathog.* **3**:e5

82. **Roth-Cross, J. K., H. Stokes, G. Chang, M. M. Chua, V. Thiel, S. R. Weiss, A. E. Gorbalenya, and S. G. Siddell.** 2009. Organ-specific attenuation of murine hepatitis virus strain A59 by replacement of catalytic residues in the putative viral cyclic phosphodiesterase ns2. *J.Virol.* **83**:3743-3753
83. **Schreiber, V., F. Dantzer, J. C. Ame, and M. G. de.** 2006. Poly(ADP-ribose): novel functions for an old molecule. *Nat.Rev.Mol.Cell Biol.* **7**:517-528
84. **Schutze, H., R. Ulferts, B. Schelle, S. Bayer, H. Granzow, B. Hoffmann, T. C. Mettenleiter, and J. Ziebuhr.** 2006. Characterization of White bream virus reveals a novel genetic cluster of nidoviruses. *J.Virol.* **80**:11598-11609
85. **Schwarz, B., E. Routledge, and S. G. Siddell.** 1990. Murine coronavirus nonstructural protein ns2 is not essential for virus replication in transformed cells. *J.Virol.* **64**:4784-4791
86. **Seybert, A., A. Hegyi, S. G. Siddell, and J. Ziebuhr.** 2000. The human coronavirus 229E superfamily 1 helicase has RNA and DNA duplex-unwinding activities with 5'-to-3' polarity. *RNA.* **6**:1056-1068
87. **Seybert, A., C. C. Posthuma, L. C. van Dinten, E. J. Snijder, A. E. Gorbalenya, and J. Ziebuhr.** 2005. A complex zinc finger controls the enzymatic activities of nidovirus helicases. *J. Virol.* **79**:696-704
88. **Seybert, A., L. C. van Dinten, E. J. Snijder, and J. Ziebuhr.** 2000. Biochemical characterization of the equine arteritis virus helicase suggests a close functional relationship between arterivirus and coronavirus helicases. *J. Virol.* **74**:9586-9593
89. **Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. Poon, Y. Guan, M. Roza-nov, W. J. Spaan, and A. E. Gorbalenya.** 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J.Mol.Biol.* **331**:991-1004
90. **Snijder, E. J., J. A. den Boon, M. C. Horzinek, and W. J. Spaan.** 1991. Comparison of the genome organization of toro- and coronaviruses: evidence for two nonhomologous RNA recombination events during Berne virus evolution. *Virology* **180**:448-452
91. **Snijder, E. J., M. C. Horzinek, and W. J. Spaan.** 1990. A 3'-coterminal nested set of independently transcribed mRNAs is generated during Berne virus replication. *J.Virol.* **64**:331-338
92. **Snijder, E. J., S. Siddell, and A. E. Gorbalenya.** 2005. The order *Nidovirales*, p. 390-404. In: B. W. J. Mahy and V. ter Meulen (eds.), *Topley and Wilson's microbiology and microbial infections*. Hodder Arnold, London.
93. **Snijder, E. J., H. van Tol, N. Roos, and K. W. Pedersen.** 2001. Non-structural proteins 2 and 3 interact to modify host cell membranes during the formation of the arterivirus replication complex. *J.Gen.Virol.* **82**:985-994
94. **Snijder, E. J., A. L. M. Wassenaar, and W. J. M. Spaan.** 1994. Proteolytic processing of the replicase ORF1a protein of equine arteritis virus. *J. Virol.* **68**:5755-5764
95. **Spaan, W., H. Delius, M. Skinner, J. Armstrong, P. Rottier, S. Smeekens, B. A. van der Zeijst, and S. G. Siddell.** 1983. Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. *EMBO J.* **2**:1839-1844
96. **Spann, K. M., J. A. Cowley, P. J. Walker, and J. G. Lester.** 1997. A yellow-head-like virus from *Penaeus monodon* cultured in Australia. *Dis.Aquat.Org.* **31**:169-179
97. **Sperry, S. M., L. Kazi, R. L. Graham, R. S. Baric, S. R. Weiss, and M. R. Denison.** 2005. Single-amino-acid substitutions in open reading frame (ORF) 1b-nsp14 and ORF 2a proteins of the coronavirus mouse hepatitis virus are attenuating in mice. *J.Virol.* **79**:3391-3400
98. **Su, D., Z. Lou, F. Sun, Y. Zhai, H. Yang, R. Zhang, A. Joachimiak, X. C. Zhang, M. Bartlam, and Z. Rao.** 2006. Dodecamer structure of severe acute respiratory syndrome coronavirus nonstructural protein nsp10. *J.Virol.* **80**:7902-7908

99. **Subbarao, K. and A. Roberts.** 2006. Is there an ideal animal model for SARS? *Trends Microbiol.* **14**:299-303
100. **te Velthuis, A. J., J. J. Arnold, C. E. Cameron, S. H. van den Worm, and E. J. Snijder.** 2010. The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. *Nucleic Acids Res.* **38**:203-214
101. **Tijms, M. A., L. C. van Dinten, A. E. Gorbalenya, and E. J. Snijder.** 2001. A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus. *Proc.Natl.Acad.Sci.U.S.A* **98**:1889-1894
102. **Tsagris, E. M., A. E. Martinez de Alba, M. Gozmanova, and K. Kalantidis.** 2008. Viroids. *Cell Microbiol.* **10**:2168-2179
103. **Unrau, P. J. and D. P. Bartel.** 1998. RNA-catalysed nucleotide synthesis. *Nature* **395**:260-263
104. **van Dinten, L. C., H. van Tol, A. E. Gorbalenya, and E. J. Snijder.** 2000. The predicted metal-binding region of the arterivirus helicase protein is involved in subgenomic mRNA synthesis, genome replication, and virion biogenesis. *J. Virol.* **74**:5213-5223
105. **Wang, K. S., Q. L. Choo, A. J. Weiner, J. H. Ou, R. C. Najarian, R. M. Thayer, G. T. Mullenbach, K. J. Denniston, J. L. Gerin, and M. Houghton.** 1986. Structure, sequence and expression of the hepatitis delta (delta) viral genome. *Nature* **323**:508-514
106. **Weiss, M., F. Steck, and M. C. Horzinek.** 1983. Purification and partial characterization of a new enveloped RNA virus (Berne virus). *J.Gen.Virol.* **64 (Pt 9)**:1849-1858
107. **Xu, Y., L. Cong, C. Chen, L. Wei, Q. Zhao, X. Xu, Y. Ma, M. Bartlam, and Z. Rao.** 2009. Crystal structures of two coronavirus ADP-ribose-1<sup>st</sup>-monophosphatases and their complexes with ADP-Ribose: a systematic structural analysis of the viral ADRP domain. *J.Virol.* **83**:1083-1092
108. **Yusupov, M. M., G. Z. Yusupova, A. Baucom, K. Lieberman, T. N. Earnest, J. H. Cate, and H. F. Noller.** 2001. Crystal structure of the ribosome at 5.5 Å resolution. *Science* **292**:883-896
109. **Zhai, Y., F. Sun, X. Li, H. Pang, X. Xu, M. Bartlam, and Z. Rao.** 2005. Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer. *Nat.Struct.Mol. Biol.* **12**:980-986
110. **Ziebuhr, J., S. Bayer, J. A. Cowley, and A. E. Gorbalenya.** 2003. The 3C-like proteinase of an invertebrate nidovirus links coronavirus and potyvirus homologs. *J.Virol.* **77**:1415-1426
111. **Ziebuhr, J., E. J. Snijder, and A. E. Gorbalenya.** 2000. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J.Gen.Virol.* **81**:853-879
112. **Zust, R., L. Cervantes-Barragan, T. Kuri, G. Blakqori, F. Weber, B. Ludewig, and V. Thiel.** 2007. Coronavirus non-structural protein 1 is a major pathogenicity factor: implications for the rational design of coronavirus vaccines. *PLoS.Pathog.* **3**:e109
113. **Zust, R., T. B. Miller, S. J. Goebel, V. Thiel, and P. S. Masters.** 2008. Genetic interactions between an essential 3' cis-acting RNA pseudoknot, replicase gene products, and the extreme 3' end of the mouse coronavirus genome. *J.Virol.* **82**:1214-1228



# Chapter 2

**Regulation of the plus-strand RNA virus  
replicative cycle: putting it together by  
taking it apart**





## INTRINSIC MULTIFUNCTIONALITY OF PLUS-STRAND RNA VIRUS GENOMES

Plus-strand RNA (+RNA) viruses with linear nonsegmented genomes and no DNA stage in their replicative cycle store their genetic information in a single molecule of mRNA polarity. This genome RNA is readily translated upon its release into the cytoplasm of the infected cell to produce a set of viral proteins which always includes an RNA-dependent RNA polymerase. Consequently, during virion assembly there is no need to package a viral “transcriptase” along with the genome, a fundamental difference between +RNA viruses and those that have minus- or double-stranded RNA genomes. The ability to express viral genes directly from the genomic RNA determines the flow of the major steps in the +RNA virus replicative cycle. Amplification of the messenger-sense genomic RNA typically follows its translation and requires the synthesis of a genome-length minus-strand RNA intermediate. Some groups of plant and animal +RNA viruses with polycistronic genomes also use the genomic RNA as a template for the generation of subgenome-length minus strands, from which subgenomic (sg) mRNAs are transcribed (reviewed in reference <sup>126</sup>). Moreover, viral genomes need to be incorporated into progeny particles in order to ensure viral propagation. This inherent multifunctionality of +RNA virus genomes might provoke “competition” for genomic RNA molecules among translation, minus-strand RNA synthesis, and encapsidation. Accordingly, +RNA viruses have evolved various control mechanisms to coordinate these processes and prevent such potential conflicts. Elucidating the molecular details of these regulatory mechanisms is technically challenging, as it entails the use of experimental approaches that separate the interconnected processes outlined above. In this manner, the viral and cellular players important for each step can be defined and characterized.

The entire replicative cycle of +RNA viruses takes place in the cytoplasm of infected cells. Specific template recognition and use in an environment where cellular mRNAs are abundant is thus required, and could be accomplished in several ways. Temporal separation of +RNA genome translation, amplification, and packaging into progeny particles would make switching between these steps well defined in time and, possibly, easier to achieve. However, the widespread use of this strategy is certainly not obvious, since all these processes occur simultaneously in cells infected with most +RNA viruses. Plus-strand RNA viruses induce profound membrane rearrangements in the cytoplasm of infected cells, which, among other functions, may allow sequestration of viral RNA and the compartmentalization of the different steps of the +RNA virus replicative cycle (reviewed in references<sup>115,125,145</sup>). Recruitment of viral and/or cellular proteins to regulatory RNA motifs in the viral genome may also determine its commitment to translation, minus-strand RNA synthesis, or encapsidation. Increase in local concentration of viral gene products could favor one of these processes in a subcellular site. The regulated *in cis* usage of genomic RNA templates could also serve as a mechanism to maximize the incorporation of translation- and replication-competent genomes in progeny virus particles by coupling those two processes with encapsidation.



This review focuses on the interplay between replication and three other genome-based processes: translation, sg mRNA synthesis, and encapsidation, in the replicative cycle of +RNA viruses. The current knowledge on the molecular mechanisms that interconnect these different steps is very fragmentary and derives mostly from studies of a few very dissimilar +RNA viruses with both mono- and polycistronic genomes, though only the latter employ the full spectrum of these processes. As an illustration of the possible strategies that +RNA viruses can employ to regulate differential use of the genomic RNA template, three relatively well-studied model systems from different +RNA virus clades are discussed below. Special emphasis has been placed on the regulatory roles of viral nonstructural proteins.

The switch between genome translation and replication has been studied extensively for the monocistronic picornavirus poliovirus (PV). Integration of data from *in vitro* and *in vivo* systems has led to a detailed model, in which the fate of the PV genome as a template for translation or minus-strand RNA synthesis is determined by the differential association of viral and host cellular proteins with regulatory RNA signals in the genomic RNA.

During replication of *Tombusviridae*, and several other groups of +RNA viruses with polycistronic genomes, subgenomic (sg) mRNAs are produced from complementary sub-genome-length minus-strand templates<sup>126,240</sup>. The determinants of viral RdRp attenuation during minus-strand synthesis therefore govern the use of genomic RNAs as templates for genome amplification or sg mRNA production, an additional aspect of +RNA genome multifunctionality. The RNA signals in particular, as well as the viral proteins involved in the “premature termination” of minus-strand RNA synthesis, have been extensively characterized for members of the *Tombusviridae*.

Recent studies on the spatial and functional coupling of genome replication and encapsidation in the monocistronic *Flaviviridae* family have revealed significant involvement of viral nonstructural proteins in coordinating these two processes and “channeling” newly synthesized viral RNA from replication to packaging. These data have effectively challenged the canonical dividing line often drawn between virion components and replicative proteins, demonstrating that apart from their “conventional” functions in ensuring viral RNA amplification, nonstructural proteins of *Flaviviridae* members are extensively involved in the late phases of the viral replicative cycle important for the generation of infectious progeny virions.

Lastly, considering the focus of this thesis on the regulatory roles of nonstructural proteins in the replication of a prototypic nidovirus, our current and mostly rudimentary knowledge on regulation of the replicative cycle in *Nidovirales* is summarized. The implications of recent progress made in research on other +RNA viruses for advancing our understanding of the molecular mechanisms governing nidovirus replication are discussed.

## CONCERTED ACTIONS OF VIRAL AND HOST CELL PROTEINS MEDIATE THE SWITCH BETWEEN TRANSLATION AND REPLICATION OF PICORNAVIRUS GENOMES

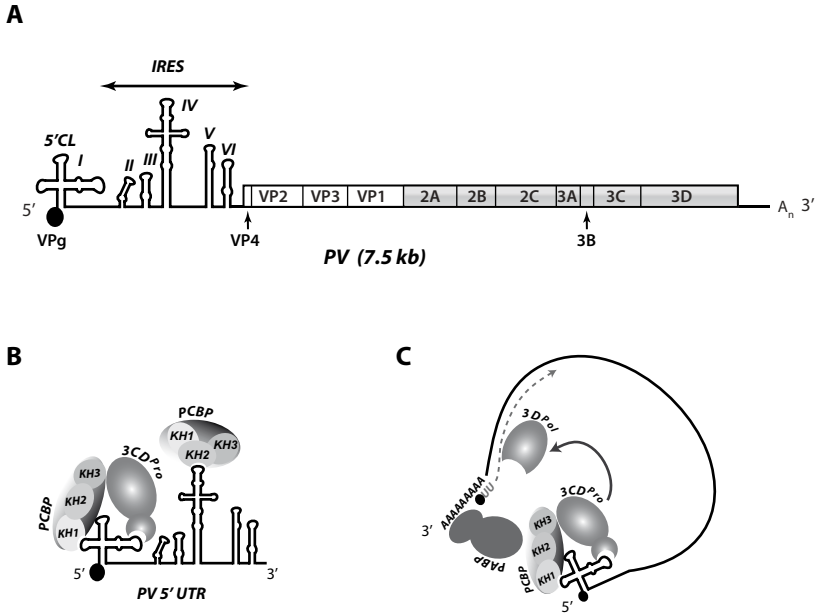
### *Picornaviridae*: an introduction

The *Picornaviridae* family of nonenveloped +RNA viruses belongs to the recently established order of *Picornavirales* and consists of 12 genera: *Enterovirus*, *Hepatovirus*, *Parvovirus*, *Kobusvirus*, *Erbovirus*, *Cardiovirus*, *Aphthovirus*, *Teschovirus*, *Sapelovirus*, *Senecavirus*, *Tremovirus*, and *Avihepatovirus*. The picornavirus group includes several important pathogens of humans and farm animals, such as PV and foot-and-mouth disease virus (FMDV), which are also the best-studied representatives of the *Enterovirus* and *Aphthovirus* genera, respectively.

Picornavirus genomes are single-stranded RNA molecules that vary between 6.7 and 8.8 kb in length and contain a single, long open reading frame (ORF), flanked by 5'- and 3'- untranslated regions (UTRs). The picornavirus genomic RNA lacks a 5' cap structure and contains a poly(A) tail at its 3' end<sup>143,249</sup>. In addition, a viral protein, VPg, is covalently linked to the 5' terminus of the genome due to its role as protein primer for genome replication<sup>54,101</sup>. The picornavirus capsid proteins are encoded in the 5' one-third of the genomic ORF, with the 3' two-thirds coding for the viral nonstructural proteins (shown for PV in Fig. 1A). Picornavirus genome translation is initiated by an internal ribosomal entry site (IRES), which is located in the 5' UTR. Translation of the genomic RNA yields a polyprotein that is co- and post-translationally cleaved by virus-encoded proteinases to generate the mature viral protein subunits and a number of processing intermediates with essential functions in viral replication. The viral proteolytic activities reside in 3C<sup>Pro</sup> and, in some viruses, also in the 3CD<sup>Pro</sup>, the 2A<sup>Pro</sup> and/or L proteins<sup>92</sup>. The picornavirus RdRp domain maps to 3D, but only the fully processed protein, and not its long-lived 3CD precursor, has this catalytic activity<sup>221</sup>. Picornavirus replication takes place exclusively in the host cell cytoplasm, where it induces extensive rearrangements of intracellular membranes into characteristic vesicular structures. Viral RNA synthesis likely occurs in association with these virus-induced vesicles<sup>19</sup> and, following encapsidation of newly synthesized genomes, the nonenveloped progeny virus particles are released upon lysis of the infected cell.

Many studies on picornavirus replication have been performed using PV as a model system. Research on the regulation of PV replication has greatly benefitted from the development of systems that reconstitutes the complete PV replicative cycle *in vitro*<sup>13,14,131</sup>. Addition of PV RNA to cell-free extracts from uninfected HeLa cells results in efficient translation of the viral polyprotein, its correct proteolytic autoprocessing, as well as *de novo* PV RNA synthesis and particle assembly<sup>131</sup>. This *in vitro* assay effectively bypasses the need for multiple rounds of replication of the incoming genomic RNA that are required in infected cells<sup>13</sup>. A similar uncoupling of translation from the initial rounds of replication was achieved in an assay for PV replication in *Xenopus* oocytes. These cells do not support

PV RNA translation, but, when a HeLa cell extract is co-injected, microinjection of PV RNA initiates a complete replication cycle, allowing the investigation of host cell factors important for viral RNA translation in the context of intact cells<sup>58</sup>. The availability of *in vitro* systems uncoupling PV translation from replication, and in particular the system based on HeLa cell extracts, has greatly facilitated mechanistic studies on these two key steps of the PV replicative cycle. Combined with complementary reverse and forward genetic approaches, *in vitro* analyses have yielded important insights into the regulation of these two interconnected processes.



**Figure 1.** A model for genome circularization and initiation of minus-strand RNA synthesis during poliovirus replication. (A) Schematic representation of the poliovirus (PV) genome. A viral peptide (VPg) encoded in the 3B region of the genome is covalently attached to the 5' end of the genomic RNA. The 5' cloverleaf (5'CL) and internal ribosomal entry site (IRES) structures in the 5' untranslated region (UTR) of the PV RNA, and the 3' poly(A) tail ( $A_n$ ) are shown. The single open reading frame (ORF) in the PV genome encodes a polyprotein, which is cleaved by proteinase activities residing in 2A, 3C and/or 3CD. The structural genes, depicted in white, are found in the 5'-proximal region of the genomic ORF and the nonstructural genes, shown in gray, are found in its 3'-proximal region. (B) Binding sites for poly(rC)-binding protein (PCBP) and 3CD<sup>pro</sup> in the 5' UTR of the PV genome (see text for details). (C) Model for circularization of the PV genome. A protein-protein bridge consisting of PCBP, 3CD<sup>pro</sup> and the poly(A)-binding protein (PABP) has been proposed to bring the viral polymerase (3D<sup>pol</sup>) in close proximity of the 3' poly(A) tail and allow for initiation of minus-strand RNA synthesis<sup>72</sup>(see text for details).

## Poliovirus genome translation as a prerequisite for its amplification: the clash between ribosomes and the RdRp

Potential coupling between genome translation and replication during PV infection was initially suggested based on results obtained using an elegant genetic approach that relied on the insertion of amber termination codons in the PV ORF. The replication capacity of PV genomes with amber mutations in the 2A<sup>Pro</sup> and 3D<sup>Pol</sup>-coding regions was assessed in amber-suppressing cell lines, and rescue of mutation-specific defects by providing a wild-type (wt) helper PV genome *in trans* was also examined<sup>144</sup>. Surprisingly, 3D<sup>Pol</sup> provided *in trans* functioned almost as efficiently as *in cis*, while no *trans*-complementation of amber nonsense mutations in 2A<sup>Pro</sup> was observed<sup>144</sup>. Consequently, this study identified a region in the 2A-coding portion of the PV genome whose translation *in cis* is required for PV replication, and the authors suggested a number of scenarios that would necessitate coupling translation to replication. An interaction of a newly synthesized viral protein with the PV RNA, for example, could be required for its establishment as a template for minus-strand synthesis, or this protein might be unable to diffuse to another template RNA due to its restricted localization, e.g. in the membrane-associated replication complex (RC).

In accordance with the latter hypothesis, formation of PV RCs has been shown to require coupled viral translation, virus-induced membrane modification, and viral RNA synthesis<sup>44</sup>. Modification of ER membranes in HeLa cells prior to PV infection, by expression of cellular or viral membrane-binding proteins, resulted in a markedly decreased capacity of these cells to support PV replication. Furthermore, previously altered membranes were not used for RC formation irrespective of their morphology<sup>44</sup>. Even the presence of membrane structures induced by translation of all PV proteins from replication-incompetent RNAs decreased the fraction of PV-susceptible cells, indicating that RNA synthesis is necessary for the formation of PV RCs and that once formed, these membrane-associated complexes do not readily exchange components.

The tight coupling between PV translation and replication implies the existence of regulatory mechanisms that prevent both processes from occurring on the same RNA template, and thus avoid potential collisions between translating ribosomes proceeding in a 5' to 3' direction, and minus-strand RNA-synthesizing complexes copying the PV genome in the opposite direction. Inhibition of PV minus-strand synthesis by translating ribosomes was indeed observed in HeLa S10 *in vitro* translation-RNA replication reactions<sup>15</sup>. Clearing PV RNA of ribosomes by addition of puromycin specifically stimulated minus-strand RNA synthesis, while addition of cycloheximide, which "freezes" ribosomes on mRNAs, had a negative effect on minus-strand production. A mechanism must therefore have evolved to coordinate initiation of minus-strand PV RNA synthesis with clearance of translating ribosomes from the viral genome during PV infection<sup>15</sup>, and differential interactions of viral and host cell proteins with RNA signals in the 5' and 3' UTRs of the PV genome have been suggested to ensure the sequential translation and replication of PV RNA.

## Formation of a ribonucleoprotein complex at the 5' UTR of poliovirus: a molecular switch?

Cis-acting RNA signals required both for PV genome translation and replication have been mapped to the UTRs of the PV genome. A *cis*-replication element (CRE), which overlaps with the 2C-coding region of the PV ORF, templates the uridylylation of VPg in a process catalyzed by 3D<sup>Pol</sup> and required for PV RNA synthesis<sup>65,66,134,138</sup>. The 5' UTR of PV is 743 nt long, and nt 134 to at least 556 forms the PV IRES that is required for allowing ribosomes to efficiently translate the PV genome without scanning it from the 5' end<sup>23</sup> (Fig.1A). While the PV IRES is crucial for viral RNA translation, it is dispensable for both minus-strand and plus-strand PV RNA synthesis<sup>139</sup>. By contrast, the 5'-proximal 90 nt of the 5' UTR fold into an RNA cloverleaf structure (5' CL) that exists only in the plus strand and is critically involved in both translation and replication of the PV genome<sup>6,7,60</sup> (Fig.1A). A 4-nucleotide insertion in stem-loop D of the 5' CL severely impaired viral RNA synthesis, and this defect could be rescued by suppressor mutations in a highly conserved amino acid stretch in the PV main protease 3C<sup>Pro</sup>, suggesting that a genetic interaction between the 5'-UTR and a 3C<sup>Pro</sup>-containing viral protein is important for PV RNA synthesis<sup>7</sup>. *In vitro* studies indeed demonstrated the formation of two 5' CL-containing ribonucleoprotein complexes in PV-infected cells - one of which composed exclusively of cellular factors and another including the viral protein 3CD<sup>Pro</sup><sup>6</sup>. Accordingly, the 4-nucleotide insertion in stem-loop D of the CL structure inhibited formation of the 3CD<sup>Pro</sup>-5' CL complex, but did not interfere with interaction of cellular proteins with the 5' CL sequence. Also, only uncleaved 3CD<sup>Pro</sup>, but not 3C<sup>Pro</sup>, 3D<sup>Pol</sup>, or a combination of these two subunits in a cleaved form, bound to the 5' CL sequence *in vitro*. These data suggested that autoproteolytic processing of 3CD<sup>Pro</sup> reduces the affinity of the resulting proteins for RNA and could thus induce disassembly of the 5' CL ribonucleoprotein complex<sup>5</sup>.

The amino acid determinants of RNA binding and proteolytic activity in the 3C<sup>Pro</sup> protein region were mapped to opposite faces of the protein surface. Mutations in 3C<sup>Pro</sup> affecting RNA binding caused specific defects in replication, without disturbing translation or stability of a PV replicon, or the proteolytic activity of 3C<sup>Pro</sup><sup>5,136</sup>. The interaction between 3CD<sup>Pro</sup> and the PV 5' UTR was not direct, however, since formation of the 3CD<sup>Pro</sup>-5' CL complex in extracts from PV-infected cells required the presence of a cellular protein<sup>5</sup>. Two widely expressed isoforms of the poly(rC)-binding protein (PCBP), PCBP1 and PCBP2, were shown to selectively bind the top of stem-loop B of the PV 5' CL structure<sup>59</sup>. PCBPs are abundant cellular proteins that contain a triple repeat of the K homology (KH) motif, present in multiple RNA-binding proteins, and can bind to 3' UTRs of cellular RNAs, thereby influencing their stability<sup>(59,160 and references therein)</sup>. PCBP1 and PCBP2 interacted with the 5' CL and stimulated 3CD<sup>Pro</sup>-CL complex formation *in vitro*. Notably, both proteins bind specifically also to stem-loop IV of the PV IRES and are required for PV translation<sup>21,22,59</sup> (Fig.1B), suggesting *trans*-acting cellular factors may be bifunctional and serve to coordinate PV translation and replication.



More detailed biochemical analyses of the interactions of PCBP2 and 3CD<sup>Pro</sup> with the 5' UTR of PV have proposed that PCBP2 can form RNA complexes with different affinities depending on the presence of 3CD<sup>Pol</sup><sup>61</sup>. The 5' CL on its own was a poor competitor of PCBP2 binding to stem-loop IV of the IRES. Simultaneous addition of 3CD<sup>Pro</sup> and 5' CL, however, effectively reduced the binding of PCBP2 to the PV IRES *in vitro*. By contrast, in the presence of PCBP, 3CD<sup>Pro</sup> associated very efficiently with the PV 5' UTR sequence *in vitro*, but not with the IRES region only<sup>61</sup>. The presence of 3CD<sup>Pro</sup> was hence proposed to promote dissociation of PCBP2 from stem-loop IV of the PV IRES and to stabilize its interaction with the 5' CL structure, providing a possible mechanism for modulating the switch between translation and replication of the PV genome.

A set of elegant experiments using the cell-free PV translation/replication assay demonstrated that active translation of a PV replicon inhibits 3D<sup>Pol</sup> elongation activity, while addition of S100 extracts from PV-infected cells negatively affects PV RNA translation. Fractionation of the inhibitory activity identified 3CD<sup>Pro</sup> as a repressor of PV RNA translation, and binding of 3CD<sup>Pro</sup> to the 5' CL was shown to be required for PV minus-strand synthesis, providing additional support for a role of the interplay between 5' CL, PCBP2 and 3CD<sup>Pol</sup> in mediating the switch between PV translation and replication<sup>60</sup>.

### **Circularization of the poliovirus genome driven by cellular proteins: a critical determinant of viral RNA synthesis?**

Poliovirus minus-strand RNA synthesis initiates within the poly(A) tail at the 3' end of the genome<sup>100</sup>. For that reason, its requirement for a 3CD<sup>Pro</sup> - 5' UTR interaction was considered unusual, until a protein-protein bridge was proposed to circularize the PV genome and thus bring the 5' and 3' RNA termini in each other's immediate proximity<sup>72</sup>. The cellular poly(A)-binding protein 1 (PABP1), an important factor controlling the stability and translation of cellular mRNAs<sup>172</sup>, interacts with the PV poly(A) tail *in vitro* with an affinity increasing as a function of the poly(A) tail length. This interaction is mediated by the N-terminal region of PABP1, while the C-terminal domain of the protein mediates association of PABP1 with PCBP2 and 3CD<sup>Pro</sup> *in vitro*, also when the two proteins are bound to the 5' CL sequence. Synergistic actions of protein-RNA and protein-protein complexes could thus mediate PV genome circularization (Fig. 1C). A mutant PABP1 with a deletion of its C-terminal domain was able to bind the PV poly(A) tail efficiently, but acted as a dominant-negative mutant when added to the cell-free PV translation/replication system. It abrogated viral RNA replication without affecting translation, providing evidence for the functional significance of the proposed genome circularization<sup>72</sup>.

Collectively, these data suggest an intricate regulatory mechanism for coordinating translation and replication of PV genomes, which relies on established intracellular protein-protein and protein-RNA interactions. Translation of input PV RNA was postulated to occur in association with cellular factors (e.g. PCBP2) until a critical concentration of viral proteins in the host cell has been reached, which is followed by shutdown of viral translation by 3CD<sup>Pro</sup>. The formation of a ribonucleoprotein complex around the 5' CL

containing PCBP2 and 3CD<sup>Pro</sup>, and subsequent interaction of this complex with PABP1 bound to the 3' poly(A), may then serve to deliver the viral RdRp precursor 3CD<sup>Pro</sup> to the initiation site for minus-strand RNA synthesis. Self-processing of 3CD<sup>Pro</sup> would then result in release of catalytically active 3D<sup>Pol</sup> and disassembly of the 5'-3' protein-RNA bridge<sup>5,60,72</sup>. According to this model, the specificity of PV minus-strand synthesis would be conferred, somewhat paradoxically, by interaction of 3CD<sup>Pro</sup> with the 5' CL structure in the genomic plus strand, since PABP1 likely associates with the poly(A) tails of cellular mRNAs as well<sup>234</sup>. As synthesis of PV proteins in infected cells continues for several hours after the onset of viral RNA synthesis<sup>246</sup>, genome translation and replication do not seem to be temporally distinct steps in the PV replicative cycle. The local concentration of 3CD<sup>Pro</sup> in a particular cellular compartment might therefore be crucial for determining the fate of PV RNAs<sup>60</sup>.

Interestingly, both PCBP2 and PABP1 are cleaved by 3C<sup>Pro</sup>/3CD<sup>Pro</sup> in PV-infected cells<sup>98,160</sup>, suggesting a mechanism for modulating the roles of these host proteins in the PV replicative cycle. PCBP proteins have three RNA-binding KH domains. KH1 was shown to bind both the 5' CL, as well as stem-loop IV of the PV IRES<sup>194</sup>, while KH2 mediates the PCBP multimerization that is important for RNA binding and translation initiation at the PV IRES<sup>16</sup>. The KH3 domain, on the other hand, interacts with SRp20, a cellular protein involved in mRNA splicing and nucleocytoplasmic trafficking. The association of PCBP2 with SRp20, which co-fractionates with ribosomal subunits, was proposed to mediate ribosome recruitment to the PV IRES<sup>160</sup>. Partial and specific cleavage of endogenous PCBP2 in the linker region between the KH2 and KH3 domains was observed in PV-infected cells from 3 hours post-infection on. *In vitro* analyses showed this reaction can be mediated by both 3C<sup>Pro</sup> and 3CD<sup>Pro</sup>, and that cleaved PCBP2 lacking the KH3 domain supported PV replication, but did not function in PV translation. These results suggest that, instead of relocating PCBP2 from the PV IRES to the 5' CL, as initially proposed<sup>61</sup>, 3CD<sup>Pro</sup> could cleave IRES-bound PCBP, thereby releasing the KH3 domain and, consequently, SRp20 from the protein complex around stem-loop IV of the IRES. Due to a subsequent lack of new initiation events, the PV genome is eventually cleared from ribosomes and becomes available for minus-strand synthesis after genome circularization.

Other results also argue against recruitment of PCBP from the PV IRES to the 5' CL by accumulating 3CD<sup>Pro</sup>. Both PCBP1 and PCBP2 co-fractionate with PV polysomes isolated from cell-free translation-replication reactions, and a mutation in 5' CL that inhibits PCBP binding also interfered with the formation and stability of nascent PV polysomes<sup>87</sup>. The association of PCBP with the PV 5' CL is therefore likely required at a very early step during PV infection, before viral proteins such as 3CD<sup>Pro</sup> could alter its properties or affinity for viral RNA structures. The partial cleavage of PABP1 observed during PV infection is also mediated by 3C<sup>Pro</sup> and occurs in the linker region between its N-terminal RNA binding portion and the C-terminal domain interacting with PCBP2<sup>98</sup>. Besides contributing to the shut-off of cap-dependent (i.e. host mRNA) translation observed during PV infection, PABP1 cleavage could also alter the ability of the cleaved protein to mediate PV genome circularization. Since not all PCBP or PABP1 molecules are cleaved during infection, the shift in the balance of PV replication and translation could occur within localized environ-



ments surrounding specific templates in which large amounts of 3CD<sup>Pro</sup> are present<sup>160</sup>. It is worthy of note that a recent report has questioned the requirement of PABP for PV RNA synthesis in cell-free extracts by observing efficient PV replication and virus production in HeLa S10 fractions depleted from a large portion of PABP<sup>203</sup>. The remaining PABP molecules, however, could still be sufficient to support efficient PV replication, given that the stoichiometry and dynamics of the protein complex proposed to circularize the PV genome<sup>72</sup> has not been examined in detail.

In conclusion, mechanistic studies of the PV replicative cycle have uncovered the existence of a complex regulatory circuit which involves the differential recruitment of viral and cellular proteins to RNA signals in the PV genome. The nature of the RNP formed may thus ultimately determine whether the messenger-sense RNA molecule will serve as a template for translation or minus-strand synthesis, and may also ensure the specificity of PV RNA amplification. Remarkably, functional coupling seems to exist not only between PV translation and replication, but also between PV replication and packaging of viral RNA into progeny particles. A requirement for active replication of PV RNAs for their subsequent encapsidation was elegantly demonstrated using a genetic approach that utilized specific inhibitors of PV RNA synthesis and virion morphogenesis<sup>146</sup>. Efficient replication of PV RNA might therefore be a prerequisite for its productive association with capsid precursors, a mechanism that would not only confer specificity to PV packaging, but also ensure that only translation- and replication-competent genomes are incorporated into progeny virus particles. A recent study has demonstrated that PV genome replication and encapsidation are discrete steps in the viral replicative cycle<sup>147</sup>, raising intriguing questions about the molecular mechanisms of coupling between these two processes.

## **LONG-DISTANCE RNA-RNA INTERACTIONS ARE MAJOR DETERMINANTS OF SUBGENOMIC RNA SYNTHESIS IN THE *TOMBUSVIRIDAE* FAMILY**

### ***Tombusviridae*: an introduction**

The *Tombusviridae*, a large family of plant viruses with monopartite +RNA genomes, includes members of the genera *Aureusvirus*, *Avenavirus*, *Carmovirus*, *Dianthovirus*, *Machlomovirus*, *Necrovirus*, *Panicovirus* and the prototype genus *Tombusvirus*. The majority of studies on tombusvirus replication and gene expression have been carried out with the prototype virus tomato bushy stunt virus (TBSV). The 4.8-kb messenger-sense TBSV genome is neither 5'-capped nor 3'-polyadenylated, and contains five ORFs. The two 5'-proximal overlapping ORFs are translated directly from the viral genome into the proteins p33 and p92, both of which are indispensable for TBSV RNA synthesis<sup>171</sup> (Fig.2A). The production of p92 requires translational read-through of the p33 amber termination codon. Consequently, p92 is N-coterminal with p33 and accumulates at a ~20-fold lower level compared to p33 during TBSV infection. Both p92 and p33 are membrane-associated<sup>182</sup> and viral RdRp signature motifs are found in the C-terminal domain of p92, which is therefore likely the catalytic



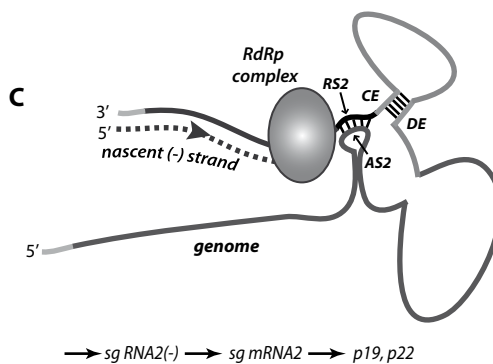
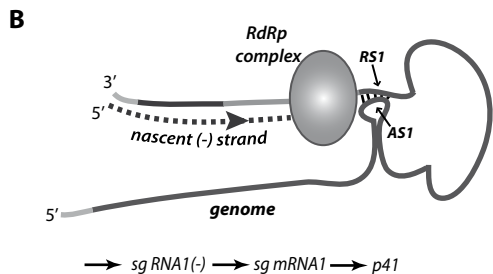
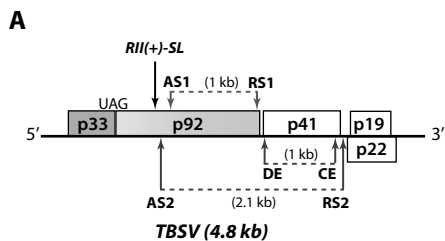
subunit for viral RNA synthesis. No known functional motifs have been identified in the sequence of the critical accessory replicative protein p33, but the N-terminal half of p33 has been shown to contain transmembrane segments and peroxisomal targeting signals that likely aid the anchoring of tombusvirus replicases to peroxisomal membranes<sup>122</sup>.

Expression of the three ORFs encoded 3'-proximally in the TBSV genome is mediated by the production of two sg mRNAs. Subgenomic mRNA1 encodes the TBSV capsid protein (CP), and sg mRNA2 serves as a template for the translation of the overlapping ORFs encoding the p22 and p19 proteins, which are involved in cell-to-cell movement and suppression of virus-induced gene silencing in infected plants. Leaky scanning, a process in which scanning ribosomes bypass the upstream start codon of the p22 ORF and initiate translation at the second start codon and in an alternative reading frame, mediates expression of p19. Although p22 is encoded 5'-proximally in sg mRNA2, translation of p19 is more efficient both *in vitro* and *in vivo*<sup>181,182</sup>. Tombusvirus sg mRNAs are presumably transcribed from 3'-truncated subgenomic minus strands that are produced by premature termination of minus-strand RNA synthesis on the genomic RNA template by the tombusvirus RdRp complex<sup>240</sup>.

### Regulatory RNA elements in the tombusvirus replicative cycle

Higher-order RNA structures and long-distance RNA-RNA interactions have been implicated in the orchestration of virtually all major steps in the tombusvirus replicative cycle, including replicase assembly, genome replication and translation, and subgenomic mRNA production. The cap- and poly(A)-independent translation of the TBSV genome is facilitated by a 5'-3' RNA-RNA base-pairing interaction that does not require protein factors and effectively circularizes the molecule<sup>50,51</sup>. Disruption of the 5'-3' interaction by the RdRp complex during initiation of minus-strand synthesis at the genomic 3' end is thus an attractive candidate switch between genome translation and RNA replication, since it would inhibit translation and clear the RdRp path of oncoming ribosomes<sup>50</sup>.

The core promoters for TBSV plus-strand and minus-strand synthesis are found at 5' and 3' termini of viral RNA molecules, but other important elements are located in protein-coding regions and can function by modulating promoter activity or influencing replicase assembly, template recruitment, or primer synthesis. One such internal RNA element, RII(+), is found in the p92-coding sequence and is required for TBSV replication<sup>132</sup> (Fig.2A). The core RII region folds into an extended hairpin structure (RII(+)-SL) that was proposed to direct the recruitment of the TBSV genome from translation to replication. Accordingly, selective binding of TBSV p33 protein to the RII sequence was demonstrated *in vitro*. Interestingly, p92 could not efficiently recognize the RII element *in vitro*, implying that template selectivity is not a function of the TBSV RdRp<sup>166</sup>. The proposed mechanism of recruitment of a RII(+)-SL -p33 complex to membranes via the hydrophobic N-terminal domain of p33 would allow the specific retention and/or amplification of the TBSV genome in RCs, since TBSV sg mRNAs lack the RII(+)-SL sequence. The RII(+)-SL element has also been suggested to assist viral replicase assembly<sup>153,243</sup>.



**Figure 2.** Regulation of subgenomic mRNA synthesis by a network of RNA signals in tobusviruses. (A) The organization of the polycistronic genome of tomato bushy stunt virus (TBSV) is depicted schematically. The viral replicase proteins, p33 and p92, are amino-coterminal and are expressed from the viral genome, the latter by read-through of the p33 amber termination codon. The ORFs located 3'-proximally in the TBSV genome are expressed from two subgenomic (sg) mRNAs. The relative positions of the RII(+)-SL RNA motif, important for TBSV replicase assembly, and RNA elements involved in long-distance interactions mediating sg mRNA synthesis are indicated (see text for details). (B-C) Models for the synthesis of TBSV sg mRNA1 (B) and sg mRNA2 (C). Long-range RNA-RNA interactions constitute "RNA attenuation signals" that mediate termination of minus-strand RNA synthesis by the TBSV RdRp complex. The resulting subgenome-length minus strands serve as templates for the synthesis of sg mRNAs. The AS1, AS2 and RS1 sequences are found in the p92-coding region of the TBSV genome. The CE and DE sequences are located in the p41 gene, while the RS2 motif is present in an intergenic region. The 5' and 3' UTRs of the TBSV genome are shown in light gray.

### An intricate network of RNA signals guides the production of tobusvirus subgenome-length RNAs

A number of well-characterized long-range intragenomic RNA-RNA interactions are critical for the generation of the subgenome-length minus-strand templates for sg mRNA production. These associations presumably induce premature termination of minus-strand synthesis by the tobusvirus RdRp complex<sup>28,29,104,240</sup>. TBSV sg mRNA1 accumulation is mediated by base-pairing between a 7-nt long "receptor sequence" (RS1), found just 3 nt upstream of the sg mRNA1 initiation site, and a complementary "activator sequence" (AS1) located ~1000 nt upstream, in the loop of a hairpin<sup>29</sup> (Fig. 2, A and B). This interaction is formed in the genomic plus strand only, and disturbing it decreases both plus- and minus-strand sgRNA1 levels. Likewise, activation of sg mRNA2 production requires two sets of interactions between noncontiguous RNA elements<sup>28,104,252</sup>. A portion

of a distal element (DE) located ~1100 nt upstream of the sg mRNA2 start site must base-pair with part of a core element (CE) located immediately 5' to the start site for efficient sg mRNA2 synthesis to occur<sup>252</sup>. The DE-CE interaction, similar to AS1-RS1 base-pairing, also occurs in the plus-strand genomic RNA, where it likely serves to reposition the sub-elements mediating a second long-range interaction between RNA elements spanning ~2100 nt<sup>28</sup> (Fig. 2, A and C). A region in the p92-coding sequence just upstream of AS1, termed AS2 and, like AS1, found in the loop of a hairpin structure, is complementary to and must base-pair with a subsection of CE termed RS2 for efficient sg mRNA2 synthesis to take place<sup>104</sup>.

Interestingly, despite the positional coupling of AS1 and AS2 in the TBSV genome, these elements do not seem to operate as a functional unit, given that sg mRNA1 accumulation is not affected by mutations decreasing sg mRNA2 levels<sup>104</sup>. The 5'-proximal location of AS1 and AS2, however, allows the genomic RNA to maintain strict control over the launching of primary sg mRNA transcripts. Since both AS1 and AS2 elements reside in the ORF of p92, their function would be negatively affected by translation, which might have to be downregulated in order for efficient sg mRNA synthesis to occur. Interestingly, the AS2-RS2 interaction could be functionally replaced by a local hairpin structure in the genomic plus strand, but this delayed accumulation of all viral RNAs and reduced the relative levels of sg mRNA2. A hairpin RNA element engineered into a TBSV defective interfering (DI) RNA could also mediate the production of a DI RNA-derived subgenomic mRNA in a sequence-independent manner. However, a minimum stability of the hairpin stem was required for RdRp attenuation, and at least one cryptic transcriptional element was found to be present in the TBSV genome. Its activity appeared to be downregulated by destabilization of the hairpin stem, likely to avoid production of superfluous and possibly deleterious viral RNAs<sup>104</sup>.

Recognition of RNA-based signals by protein factors in the tombusvirus RdRp complex is, without a doubt, a key determinant of their regulatory activities. Furthermore, components of the viral RNA-synthesizing complex might also modulate formation or disruption of these RNA-RNA interactions to ensure the temporal and quantitative regulation of the different steps they mediate. Accordingly, C-terminal modifications of p92 specifically disrupted TBSV sg mRNA production without affecting genome amplification, suggesting these two processes can be effectively uncoupled via the viral RdRp<sup>244</sup>. Due to the overlap of RNA signals that mediate sg mRNA1 production with the region coding for the C-terminal part of p92, a TBSV clone with translationally silent mutations in RS1 was engineered first. As a result, sg mRNA1 synthesis was inactivated, and therefore only the effects of RdRp sequence modifications on sg mRNA2 accumulation could be studied. Reverse genetics analyses of the C-terminal domain of p92 revealed a progressive decrease in sg mRNA accumulation levels upon removal of up to 4 residues, while genomic RNA levels were not affected. Deletion of six or more C-terminal residues, however, resulted in nonviable phenotypes<sup>244</sup>. The TBSV RdRp was found to function primarily at an early step during premature termination of TBSV minus-strand synthesis, which is important for the accurate and efficient production of subgenome-length minus-strand templates. The

ability of C-terminally truncated p92 to recognize different RNA attenuation signals, and their base-paired components in particular, seemed compromised in the mutant RdRps. These results identified the tombusvirus RdRp-containing replicase subunit as the key protein factor mediating recognition of attenuating RNA signals that control subgenome-length minus-strand production, arguing against differences in protein composition of RNA-synthesizing complexes producing genome-length and subgenome-length minus strands, in contrast to what has been proposed for arteriviruses<sup>212</sup>.

### **Regulation of the relative abundance of tombusvirus mRNA species**

Tombusvirus sg mRNAs serve as templates for translation of the viral structural proteins. Subgenomic mRNA accumulation levels in tombusvirus-infected protoplasts are roughly proportional to the amounts of their translation products, suggesting transcriptional control as a mechanism to regulate viral protein expression levels<sup>233</sup>. Since the sg mRNA promoters and the genome plus-strand promoter are similar linear sequence elements<sup>104</sup>, tombusvirus mRNA abundance is likely determined by the accumulation levels of minus-strand templates, and thus (also) by the stability of the different RNA signals that mediate attenuation. Extending the complementarity region between AS2 and RS2 sequences, however, did not result in increased sg mRNA2 levels<sup>104</sup>.

The critical role of higher-order RNA structure in attenuation of minus-strand synthesis by the tombusvirus RdRp, on the other hand, was elegantly demonstrated by introduction of theophylline-binding RNA aptamers as the RNA attenuation components for sg RNA2 synthesis<sup>233</sup>. Addition of the ligand for the aptamer stabilized its structure and induced efficient sg mRNA2 production, arguing against a fundamental requirement for a particular protein factor to specifically bind to the attenuation signal in order for it to function, or for induction of a conformational switch in the RdRp. These data imply that the activity of tombusvirus RNA attenuation signals is primarily structure-independent, and they might instead act as physical barriers impeding the progress of the RdRp on the genomic plus-strand template. In line with this hypothesis, destabilizing the stem of the AS2-containing hairpin structure stimulated sg mRNA2 production, suggesting a potential mechanism for transcriptional regulation of p19 and p22 expression<sup>105</sup>. Suboptimal attenuation of minus-strand synthesis in the sg mRNA2 promoter region might serve to ensure that a sufficient number of RNA-synthesizing complexes will reach the sg mRNA1 promoter region located upstream, and therefore sufficient mRNA templates for CP translation will be produced. The lack of increased sg mRNA1 accumulation in the absence of sg mRNA2 production, however, argues against this hypothesis<sup>104</sup>.

Alternatively, the suboptimal levels of sg mRNA2 might be related to specific properties of the proteins they encode, as has been suggested for cucumber leaf spot virus (CLSV), member of the genus *Aureusvirus*. Subgenomic mRNA2 of aureusviruses encodes two proteins - p17 (suppressor of gene silencing) and p27 (movement protein). The suppressor protein of CLSV has been shown to induce extensive tissue necrosis during plant infection, which can adversely affect optimal virus spread. Accumulation levels of



sg mRNA2 in aureusvirus-infected cells are significantly lower than those of sg mRNA1, suggesting a specific downregulation of the abundance of this transcript to minimize suppressor protein-associated toxicity to the host. Functional counterparts of the TBSV AS2 and RS2, but not of the DE and CE sequences were identified in CLSV<sup>245</sup>. The CLSV AS2-RS2 interaction was extremely sensitive to destabilization. Replacements stabilizing the higher-order component of the CLSV sg mRNA2 attenuation signal or optimizing the sg mRNA2 promoter sequence resulted in elevated sg mRNA2 levels, suggesting that both the AS2-RS2 interaction and the sg mRNA2 promoter are operating suboptimally in CSLV and thus contribute to the consistently lower levels of sg mRNA2 observed.

Experimental evidence suggests that tombusvirus replicative proteins are also involved in controlling the relative abundance of viral mRNAs. Mutations in the RS1 sequence that alter the C-terminal residues of p92 had a negative impact on the accumulation of both sg mRNA1 and sg mRNA2. While introduction of complementary substitutions in AS1 partially restored accumulation of sg mRNA1, no restoration of sg mRNA2 levels was observed<sup>29</sup>, suggesting that the C-terminal domain of the TBSV RdRp is less critical for sg mRNA1 production. A function in modulating the relative abundance of viral mRNAs was also proposed for an arginine/proline-rich RNA-binding motif (RPR), found in the sequence encoding the common part of p33 and p92 of the carmovirus cucumber necrosis virus (CNV)<sup>152</sup>. Reverse genetics revealed that certain mutations in the RPR motif considerably altered the molar ratios at which CNV RNAs accumulate relative to each other, though this was mostly observed for mutants in which overall viral RNA accumulation was severely decreased<sup>152</sup>. Mutagenesis of a phosphorylation site adjacent to the p33/p92 RPR motif also had an effect on the molar ratios of sg mRNA1 and sg mRNA2 in CNV-infected cells<sup>189</sup>. Using a two-component system, in which p92 is expressed from the CNV genomic RNA and p33 from a CNV DI RNA, an essential role for the RPR motif of p33 in CNV replication was demonstrated. In the same system, replacements in this motif had little effect on p92 function, possibly due to the presence of two additional RNA-binding domains in its central and C-terminal portion<sup>169</sup>. This experimental set-up unfortunately did not allow the definite assignment of a role in controlling the relative abundance of CNV mRNAs to either p33 or p92, since only a single RNA molecule was amplified.

### **Host proteins involved in tombusvirus RNA synthesis**

Reconstitution of key steps of tombusvirus replication in tractable model hosts and cell-free *in vitro* systems have provided valuable tools for studying the roles of viral and host replication proteins in tombusvirus RNA synthesis. Highly active, template-dependent CNV replicase complexes have been purified from *Saccharomyces cerevisiae*, taking advantage of the ability of p92 and p33 proteins to direct the amplification of a cognate DI RNA template when co-expressed in this host<sup>153</sup>. TBSV replication has also been reconstituted in yeast cells and, notably, also in cell-free yeast extracts<sup>151,164</sup>. TBSV replication was shown to occur in a membranous compartment in the cell extract, where the viral RNA is protected from nuclease degradation<sup>164</sup>. Proteomics analysis of the highly

purified tombusvirus replicase complex revealed the presence of a number of host proteins, among which the yeast orthologs of glyceraldehyde-3-phosphate dehydrogenase and heat-shock protein 70 - Tdh2p/3p and Ssa1p/2p<sup>165,232</sup>. The presence of Ssa1p/2p is required for the *in vitro* assembly of a functional TBSV replicase, possibly to facilitate re-folding of replicase components required for the formation of an active RC<sup>165</sup>. By contrast, a specific and significant reduction in TBSV plus-strand RNA accumulation was observed in yeast cells lacking Tdh2p and its ortholog Tdh3p. Consequently, these proteins were implicated in promoting efficient use of minus-strand genomic RNA as a template by its retention in the RC<sup>232</sup>. Since this “RC retention signal” is only present in genome-length minus-strand RNA, it would be interesting to determine whether the subgenome-length minus-strand templates for the synthesis of sg mRNA1 and sgmRNA2 are also retained in active RCs by a Tdh2p/3p-mediated mechanism. Despite that a number of host proteins affecting tombusvirus RNA synthesis and recombination have been identified<sup>83,185</sup>, no host factors have been implicated in promoting tombusvirus sg mRNA synthesis so far. Active tombusvirus replicase complexes purified from yeast or plant cells do not mediate sg mRNA production *in vitro*, however, suggesting their functionality might be perturbed by the isolation procedure or the fact that important cofactors for this process are missing<sup>244</sup>.

In summary, local and long-distance RNA-RNA interactions and higher-order RNA structures are extensively involved in coordinating the use of tombusvirus genome as a template for the production of genome-length and subgenome-length minus strands. These interactions are mainly thought to mechanistically promote termination of minus-strand synthesis by the tombusvirus RdRp at sg promoter regions, and have been fine-tuned to guarantee the optimal accumulation levels of tombusvirus sg mRNAs and their translation products<sup>233,245</sup>. How is formation of the mutually interfering RNA attenuation signals regulated, however (Fig. 2, B and C), remains a question that is largely unexplored. Protein factors may assist in bringing together RNA motifs separated by large distances in the tombusvirus genome, and both viral replicase subunits have been implicated in the specific recognition of the RNA signals critical for the production of subgenome-length minus strands. Taking into account the small coding capacity of the tombusvirus genome and the significant involvement of host cell proteins in crucial steps of tombusvirus replication outlined above, a role for host proteins in the temporal and/or qualitative modulation of tombusvirus “RNA attenuation signal” formation certainly merits further investigation.



## NONSTRUCTURAL PROTEINS ARE KEY REGULATORS OF FLAVIVIRUS GENOME ENCAPSIDATION AND VIRION ASSEMBLY

### *Flaviviridae*: an introduction

The *Flaviviridae* family is a group of small, enveloped +RNA viruses that consists of three genetically distinct genera - *Flavivirus*, *Pestivirus* and *Hepacivirus*, with members of the latter two genera being more closely related to each other than to *Flavivirus* representatives<sup>123</sup>. The *Flavivirus* genus comprises more than 80 members and includes a number of vector-borne viruses, such as yellow fever virus (YFV), dengue virus (DENV) and West Nile virus, which are associated with severe human diseases. The prototype of the *Hepacivirus* genus, hepatitis C virus (HCV), is a blood-borne human pathogen that frequently establishes chronic infection and can lead to significant liver damage, while members of the genus *Pestivirus*, such as bovine viral diarrhea virus (BVDV) and classical swine fever virus (CSFV), are economically important animal pathogens.

The single-stranded RNA genome of *Flaviviridae* members varies between 9.5 and 12.5 kb in length. It contains a single ORF, with the viral structural proteins encoded in the 5' one-third and the nonstructural proteins in the 3' two-thirds (Fig. 3). The genomic RNA of *Flavivirus* representatives contains a 5' cap structure, but lacks a 3' poly(A) tail and the genomic 3' UTRs contains a number of highly conserved RNA secondary structure elements that are important determinants of genome translation and replication. HCV and pestivirus genome translation, on the other hand, is mediated by an IRES found in the 5' UTR of the viral genome, and the 3' UTR is much shorter and less structured<sup>210</sup>.

The viral proteins are translated from the flavivirus genome in the form of a polyprotein that is processed by cellular and viral proteinases to yield the individual viral gene products. Proteolysis of the flavivirus structural proteins, which consist of the capsid protein (C; called core in HCV) and envelope glycoproteins, is mediated predominantly by host cell peptidases. The nonstructural proteins are processed by virus-encoded (auto)proteinases and form the viral replication complexes, which mediate viral RNA synthesis in close association with modified membranes in the host cell cytoplasm (reviewed in<sup>107,133</sup>).

The principal functions of +RNA virus nonstructural proteins are related to modification of the intracellular environment to accommodate viral RCs and orchestration of the synthesis of progeny viral genomes. Increasing evidence, however, suggests the existence of a high degree of "cross talk" between nonstructural and structural viral gene products that is critical for the late stages of the replicative cycle, such as selective genome encapsidation, infectious particle assembly, and release. Much of this evidence has been obtained for nonstructural proteins of *Flaviviridae*. The key enzymes required for genome expression and replication - the viral protease and helicase (residing in NS3) and the RdRp (found in NS5) are conserved in the entire *Flaviviridae* family (105). The *Hepacivirus* and *Pestivirus* viruses share an additional conserved domain - a protease which maps to NS2<sup>99</sup>. The remaining replicase domains of *Flavivirus* and *Hepacivirus* members are not evolutionary related, and the molecular details of virus particle formation in these two

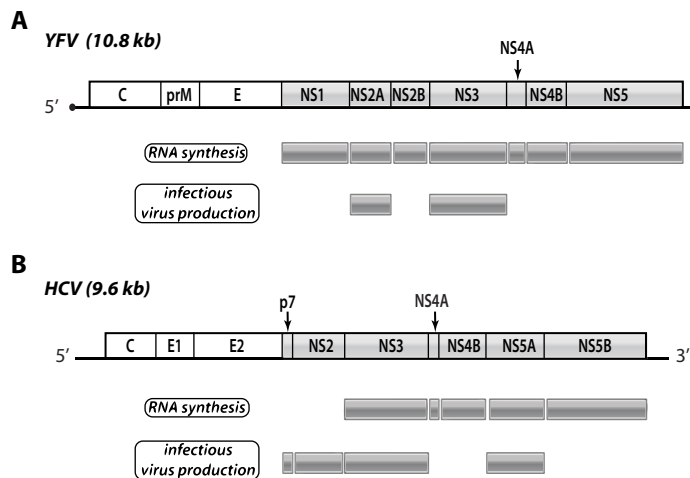


genera also differ. Therefore, the nonstructural protein determinants of virion biogenesis in *Flavivirus* and *Hepacivirus* representatives are discussed separately below.

### Enzymatic and regulatory activities of *Flavivirus* nonstructural proteins in virion production

The *Flavivirus* polyprotein organization is schematically depicted in Fig. 3A. The N-terminal third of NS3 contains a trypsin-like serine protease domain, the folding and protease activity of which require the noncovalent association with a hydrophilic domain in NS2B<sup>48,52</sup>. The C-terminal portion of NS3 has helicase (Hel) and NTPase activity<sup>239</sup>, while NS5 has both 2'-O-methyltransferase and RdRp activities<sup>46,204</sup>. NS2A, NS2B, NS4A and NS4B are highly hydrophobic proteins with small hydrophilic segments and poorly understood roles in the flavivirus replicative cycle. Kunjin (KUN) virus NS2A co-localizes with dsRNA, NS3 and NS5<sup>117</sup>, and also binds the KUN 3' UTR *in vitro*, while both NS2A and NS4A has been implicated in inducing the membrane alterations with which flavivirus RCs are associated in infected cells<sup>102,124</sup>.

Flavivirus progeny particles are formed by envelopment of nucleoprotein complexes by ER-derived membranes that contain the viral glycoproteins, and the newly formed virions are transported to the cell surface via the secretory pathway<sup>118</sup>. Subgenomic flavivirus replicons lacking the structural proteins replicate efficiently and can be packaged by expression of the structural proteins *in trans*<sup>89</sup>. Packaging signals, however, have not



**Figure 3.** Multiple nonstructural proteins of *Flaviviridae* are involved in particle assembly. (A) Schematic representation of the genomes of yellow fever virus (YFV, genus *Flavivirus*, panel A) and hepatitis C virus (HCV, genus *Hepacivirus*, panel B). The YFV genomic RNA carries a cap structure at its 5' end, whereas HCV employs a 5'-proximal IRES element. The single ORF in the YFV and HCV genomes is translated into a polyprotein, which is proteolytically processed by host cell and virus-encoded proteinases. Structural proteins are shown in white and nonstructural proteins are shown in gray. The nonstructural proteins required for viral RNA synthesis and implicated in infectious virus production are highlighted.

been identified in flavivirus genomic RNAs to date, and the lack of an ordered nucleocapsid interior that is characteristically observed in electron microscopy reconstitutions of flavivirus particles has led to the suggestion that the C protein interacts electrostatically with the viral RNA to aggregate and condense it into particles. Moreover, co-expression of flavivirus prM and E is sufficient for secretion of "subviral particles" that mimic the major properties of the mature virion envelope<sup>178</sup>. These observations suggest that flavivirus membrane glycoproteins play major roles in orchestrating virion morphogenesis, although the determinants of the highly selective packaging of flavivirus RNAs<sup>89</sup> remain poorly understood.

The majority of regulatory functions of nonstructural proteins in virion morphogenesis described to date are largely independent of their known enzymatic activities. A direct enzymatic role of flavivirus nonstructural proteins in the regulation of the process of virion formation has been shown only for the main protease from the *Flavivirus* genus, NS3, and its cofactor, NS2B. An NS3-NS2B complex mediates the proteolytic release of the viral C protein from its membrane anchor, which occurs on the cytoplasmic side of the ER membrane. This cleavage is necessary for the subsequent processing of the viral envelope protein, prM, by the host enzyme signal peptidase in the ER lumen. Sequential processing of two major virion components initiated by a viral protease was proposed to delay virus budding until sufficient genomic RNA has accumulated in infected cells, and to reduce secretion of immunogenic subviral particles<sup>110</sup>.

Recent biochemical and imaging studies have suggested the existence of virus-induced microenvironments in infected cells that allow the coupling of genome amplification to its incorporation into progeny particles via specific interactions among structural and nonstructural proteins. Ultrastructural studies of DENV-infected cells have indeed suggested a possible topological link between DENV replication and assembly sites<sup>238</sup>. DENV replication induces various modifications of host cell membranes, such as convoluted membranes (CMs), thought to be the sites of viral polyprotein processing, and vesicle packets, which are likely associated with viral RNA synthesis<sup>116</sup>. Examination of DENV-infected cells by transmission electron microscopy showed that virus particles could be found in the ER lumen and occasionally in the immediate proximity of vesicles and CMs. Three-dimensional reconstruction of cellular volumes containing DENV-induced membrane alterations revealed the presence of a continuous ER-derived membrane network containing the virus-induced vesicles and CMs. Moreover, the vesicles likely represented invaginations of the ER membrane that are connected to the cytosol via neck-like channels, which in turn could serve as exit sites for the transport of newly-synthesized viral RNA to the cytosol for translation or encapsidation<sup>238</sup>. Accordingly, ER cisternae containing virus particles could be found in association with this network and in close proximity to the pores of replication vesicles, suggesting spatial coupling of DENV genome replication and its subsequent incorporation into progeny particles.

Coupling of genome replication to packaging has also been proposed for Kunjin (KUN) virus based on data obtained with a replicon system, in which KUN replicon RNAs are produced by RNA polymerase II-driven system from plasmids containing their cDNA

sequence<sup>88</sup>. Viral replication is thus effectively uncoupled from protein production and genome encapsidation, allowing the investigation of the importance of KUN replicon RNA translation and replication for its incorporation into virus particles. Using a construct with a deletion in the viral RdRp active site<sup>227</sup>, it was shown that replication-incompetent KUN virus RNA was only packaged into secreted virions when it was produced in cells where it could be replicated via complementation<sup>88</sup>. In addition to establishing replication of KUN virus RNA as a prerequisite for its incorporation into progeny virions, these data also suggest that the establishment of virus-induced membrane-associated replication sites might be essential for its relocation to virion assembly sites.

Consistent with the spatial coupling of flavivirus replication and particle formation, a number of reports have described essential roles for flavivirus NS2A and NS3 in virion morphogenesis. NS2A of YFV is found in two forms in infected cells– a full-length protein (224 aa) and a C-terminally truncated form (NS2A $\alpha$ ; 190 aa) resulting from partial processing by the main viral protease at K190. A K190S substitution in the full-length YFV clone blocked NS2A $\alpha$  cleavage and selectively inhibited production of infectious particles in the absence of defects in viral RNA synthesis<sup>96,141</sup>. Detailed characterization of NS2A $\alpha$  cleavage site mutants revealed that secretion of nucleocapsid-containing virions, but not of subviral particles, was impaired. Expression of NS2A *in trans* from an alphavirus replicon could complement the defect in infectious virus production of the K190S mutant<sup>96</sup>. Interestingly, plaque-forming viruses recovered after electroporation of an NS2A $\alpha$  cleavage site mutant defective in infectious progeny release contained compensatory second-site mutations in the Hel domain of NS3 (D343 to V, A or G) that restored infectivity to near wildtype levels without rescuing the NS2A $\alpha$  cleavage. Taken together, these data provide genetic evidence for an interaction of NS2A with the Hel domain of NS3 that is critical for infectious virus productions. They also imply that a downstream effect of the NS2A internal cleavage which could be also brought up by a mutation in NS3, rather than NS2A cleavage *per se*, is important for this process.

An essential role in virus assembly has also been demonstrated for NS2A from KUN virus<sup>108</sup>. KUN RNA transcribed from a full-length cDNA clone carrying a NS2A I59N mutation did not support virus particle release<sup>108</sup>. Analogous to the defects in virus production associated with the K190S mutation in YFV NS2A, secretion of KUN subviral particles was not affected by the I59N substitution<sup>102</sup>. Complementation analyses was used to demonstrate that KUN RNAs with the I59N mutations could be packaged into virions by the KUN virus structural proteins produced *in cis* from the same RNA molecule, but the presence of wild-type NS2A produced either *in cis* or *in trans* was absolutely essential for packaging RNA into virions or virus-like particles<sup>108</sup>. Immunofluorescence and electron microscopy analyses revealed an intriguing phenotype of the NS2A-I59N replicon-expressing cells. Virus-induced membrane structures normally present in cells expressing wild-type replicon RNA were essentially lacking in the I59N mutant, while a compensatory mutation in NS2A restored their induction to a level similar to those observed during wt KUN virus replication. These results not only implicated NS2A in the biogenesis of virus-induced membrane structures, but lent further support to the hypothesis that these structures

play a critical role in flavivirus assembly<sup>88,238</sup>. Combining the data from biochemical and imaging studies thus favors a model in which flavivirus NS2A is directly involved in transporting RNA from sites of RNA replication to adjacent sites of virion assembly that could very well be part of the same virus-induced membrane network<sup>102</sup>.

The role of flavivirus NS3 in the production of infectious particles is still poorly understood, although a recent report suggests it does not depend on the protein's helicase/NTPase function<sup>158</sup>. Mutations of D343 in the Hel domain of NS3 rescued a block in infectious virus production caused by a substitution in NS2A<sup>96</sup>, and an alanine replacement of W349, neighboring D343 in a flexible solvent-accessible loop in the NS3 Hel domain structure, resulted in a specific block of infectious virus production<sup>158</sup>. Similarly to the defects in assembly upon NS2A mutagenesis, the W349A mutation in NS3 did not hamper the release of subviral particles, and wild-type NS3 produced from a YFV replicon could complement, although very inefficiently, the virion assembly defect of the W349A mutant. By contrast, NS3 is required *in cis* for KUN virion assembly<sup>109,163</sup>. Using a dual replicon complementation system, Liu and colleagues showed that defects in KUN RNA replication upon deletion of the Hel/NTPase domain-coding sequence can be complemented *in trans*, while the N-terminal 178 codons (coding for the protease domain) are required *in cis* for KUN RNA synthesis. A complete block of virus production occurring at the stage of virus assembly was associated with both deletions in NS3, however, and could not be complemented *in trans*<sup>109</sup>. RNA structures in the KUN virus NS3 coding sequence do not seem to play a role in RNA replication and virion morphogenesis, effectively ruling out the presence of RNA packaging signals in this genomic region<sup>163</sup>. By contrast, NS3 translation *in cis* is strictly required for subsequent KUN genome encapsidation. Co-translational association of NS3 with a progeny viral genome exiting the RC might thus be one of the determinants for specific packaging of replication- and translation-competent KUN RNAs. The less strict requirement for YFV NS3 *in cis*<sup>86,158</sup>, on the other hand, suggests that important differences in the virion assembly process might exist among members of the *Flavivirus* genus.

### **Nonstructural proteins of hepatitis C virus coordinate the assembly of infectious particles**

Over 130 million individuals are estimated to be persistently infected with HCV, the prototype (and so far sole member) of the *Hepacivirus* genus. Based on genetic differences, naturally occurring HCV variants are classified into six major genotypes (1-6) with several subtypes within each genotype (represented by letters). HCV-related liver disease frequently progresses to cirrhosis that can lead to hepatocellular carcinoma. Development of effective therapies and a vaccine against this major human pathogen were initially greatly hindered by the inability of HCV to grow efficiently in cell culture. Subgenomic RNA replicons encoding HCV NS3 to NS5B have been adapted for efficient RNA replication in the human hepatoma cell line Huh-7<sup>112</sup>, but for a long time studies of the entire HCV replicative cycle were not feasible, because full-length genomes containing cell

culture-adaptive mutations did not produce infectious virus particles in cell culture<sup>20</sup>. This major obstacle was overcome by construction of full-length chimeric genomes in which the HCV “replicase” (NS3-NS5B) and UTRs were derived from the unique genotype 2a patient isolate JFH1, which can replicate efficiently in cell culture without adaptive mutations<sup>106,161</sup>. The full-length chimeric HCV RNAs, as well as that of the complete JFH-1 genome<sup>231,253</sup>, are competent for RNA replication and also direct the efficient production of infectious progeny particles. The availability of this experimental system has greatly facilitated research on the mechanisms underlying HCV entry, as well as virion assembly and egress, in the context of the full HCV replicative cycle.

The gene organization within the HCV polyprotein (Fig. 3B) is similar to that of members of the *Pestivirus* genus, which will be discussed only briefly here. NS3 to NS5B are essential for and sufficient to support HCV RNA replication. NS3 has Hel/NTPase activities and, together with its key cofactor NS4A, comprises the main viral proteinase. NS5B has RdRp activity<sup>111</sup>, while NS4B is a hydrophobic protein with numerous predicted membrane-spanning domains that is implicated in the formation of the “membranous web”, a virus-induced membrane structure derived from the host cell ER and presumably harboring the HCV RCs<sup>45,70</sup>. NS2 and p7 also contain hydrophobic regions, but have no documented role in RNA replication and are not incorporated into HCV particles to detectable levels. Ion channel activity has been demonstrated for p7<sup>30,71</sup>, while NS2 is an integral membrane protein that is tightly associated with intracellular membranes via a highly hydrophobic N-terminal domain. The cytosolic portion of NS2 contains a cysteine protease domain that, in association with the N-terminal portion of NS3, mediates the autoproteolytic processing of the NS2-NS3 junction<sup>113</sup>. While NS2 itself is dispensable for HCV replication, processing of the NS2/NS3 site is crucial for viral RNA synthesis<sup>236</sup>. The crystal structure of the post-cleavage form of NS2 revealed a dimeric cysteine protease with two composite active sites, a peculiar feature essential for the autoproteolytic release of NS2<sup>113</sup>.

A number of recent reports suggest that, as for distantly related members of the *Flavivirus* genus, nonstructural proteins are also critically involved in the morphogenesis and secretion of HCV particles. A report analyzing the stoichiometry of HCV protein and RNA molecules estimated that only a minor fraction of the HCV nonstructural proteins (<3%) is actively involved in RNA synthesis at a given time point, and a large excess of viral proteins over viral RNAs is produced during HCV infection<sup>168</sup>. This study was carried out in a system allowing only transient replication of full-length HCV genomes with replication-enhancing mutations, which, in addition, did not support virus particle production. It would be thus necessary to address this issue in the context of the complete HCV replicative cycle, and assess the effects (if any) of active particle morphogenesis on the stoichiometry of viral proteins and RNA molecules. Notably, a recent report confirmed the hypothesis that replication-enhancing mutations in HCV replicase proteins interfere with the production of infectious HCV particles in cell culture, arguing that enhanced replication triggered by these mutations occurred at the expense of genomic RNA encapsidation and particle formation<sup>93,162</sup>.



Infectious HCV virions are likely assembled at ER membranes as high-density particles and undergo subsequent transformation to lower-density particles before or during their export to the extracellular milieu<sup>62</sup>. The lipid droplet, a cellular organelle involved in the storage of neutral lipids, plays a central role in HCV assembly<sup>128</sup>. Utilizing lipid droplets as assembly organelles is presumably linked to their contribution to the formation of very-low-density lipoproteins in hepatocytes, since HCV particles circulating in infected patients are associated with lipoproteins. Interestingly, an interaction of DENV C protein and lipid droplets that is critical for efficient virus production was recently reported<sup>173</sup>, suggesting the involvement of this ubiquitous cellular organelle in virion morphogenesis might be a common property in the *Flaviviridae* family.

Recruitment of core protein and NS5A to the lipid droplet is a major determinant of infectious HCV particle assembly<sup>128,190</sup>. NS5A is a phosphoprotein found in a hypo- and hyperphosphorylated form in HCV-infected cells<sup>205</sup> and the majority of mutations conferring enhanced replication to HCV subgenomic replicons in cell culture map to the NS5A-coding region. The protein is presumably anchored to membranes via an N-terminal  $\alpha$ -helix. Three subdomains, separated by low-complexity regions, have been indentified in NS5A so far. Domain I, located immediately downstream of the  $\alpha$ -helix, is the most conserved across HCV genotypes and binds RNA in homodimeric NS5A<sup>75,208</sup>. The region required for NS5A association with lipid droplets is also located within domain I<sup>128</sup>. The roles of the less conserved domains II and III of NS5A in the HCV replicative cycle are less well understood, although at least part of domain II was shown to be essential for HCV RNA synthesis<sup>8</sup>. Domain III is likely involved in basal phosphorylation of NS5A, and while it is dispensable for RNA replication, residues from this domain are critically involved in infectious progeny production. On the basis of these observations, NS5A was proposed to act as a phosphorylation-dependent switch between genome replication and particle assembly<sup>179,207</sup>. An assessment of the influence of deletions in domain II and domain III on the phosphorylation status of NS5A, however, revealed that deletion of residues in domain II of NS5A important for its hyperphosphorylation had little effect on HCV replication and particle assembly<sup>8</sup>. In contrast, deletions in domain III strongly interfered with virion morphogenesis. This phenotype was not related to lack of NS5A at lipid droplets, but rather to a disturbed association between NS5A and core protein on the same lipid droplet. An interaction between NS5A and core protein, mediated by a Ser cluster in the C-terminal region of NS5A, was shown to be a prerequisite for infectious particle production<sup>121</sup>. Interestingly, substitutions of Ser residues from this cluster impaired the basal phosphorylation of NS5A, suggesting this posttranslational modification might play a role in NS5A-core association, which is also seemingly required for the interaction between core and the HCV genome<sup>121</sup>. The defects in virus production associated with deletions in domain III of NS5A could be complemented *in trans* by wt NS5A expressed from a subgenomic HCV replicon, i.e. in the context of a replicase, but not when expressed on its own. Overall, NS5A seems to play a key role not only in genomic RNA amplification, but also in its recruitment from the membrane-associated RCs to lipid droplets via a core-NS5A association, implying that protein-protein rather than protein-RNA interactions are responsible for the specificity of HCV RNA encapsidation<sup>121</sup>.

In addition to NS5A, both p7, NS2 and NS3 are critically involved in the morphogenesis and release of infectious HCV particles<sup>85,202,247</sup>. p7 promotes virus assembly in a genotype-specific manner, suggesting that interactions with other viral factors are important for this function<sup>202</sup>. Also, the assembly defects of HCV genomes with certain mutations in p7 could be rescued by providing p7 *in trans*<sup>26</sup>. The NS2 protease domain is important for the protein's role in virion assembly, but both its catalytic activity and uncleaved NS2 precursor forms are nonessential<sup>84,85</sup>, in contrast to a documented critical role for uncleaved NS2-3 in the assembly or egress of the pestivirus BVDV particles<sup>1</sup>. The amino acid determinants for the protease activity and post-cleavage functions of NS2 seem to differ. C-terminal truncations of NS2 interfere both with virus assembly and RNA replication, while replacement of the C-terminal Leu residue or fusion of additional amino acids to the C-terminus of NS2 abolish infectious virus production while preserving full RNA replication capability<sup>39</sup>. In addition, results from analyses of intra- and intergenotypic chimeric HCV genomes have shown that the highest HCV infectivity titers are achieved with chimeric genomes in which the junction is positioned in the loop region connecting the two transmembrane segments of NS2<sup>161</sup>. The first hydrophobic region of NS2 is therefore also important for HCV virus production, and consequently, full-length NS2 was shown to be required for particle assembly. Structural and biochemical analyses have identified a Gly residue located in the N-terminal transmembrane region of NS2 that is critical for virion assembly and might be involved in intra-membrane protein-protein interactions<sup>84</sup>. Moreover, Arg-to-Lys substitutions in the transmembrane regions of NS2 reduced infectious particle production by decreasing the specific infectivity of secreted virions<sup>237</sup>.

NS2, when expressed on its own, is rapidly degraded via the proteasome and NS2 stability was proposed to be regulated by phosphorylation at S168<sup>55</sup>. This highly conserved residue, however, does not seem to affect NS2 stability in the context of a replicating HCV genome, but S168A and S168G replacements severely decrease infectious virus production<sup>84,248</sup>. The S168G mutation did not prevent the assembly of high-density virus-like particles, suggesting NS2 mediates the "maturation" of noninfectious intracellular HCV particles to infectious particles that can be released from infected cells. NS2 might thus mediate a rearrangement of viral envelope proteins subsequent to particle assembly and be required for infectivity. Accordingly, NS2 strongly co-localized with E2 and partially with NS5A, but not with core protein in infected cells, and substitutions of S168 do not alter the NS5A-core association at lipid droplets<sup>248</sup>. Pseudorevertant analysis mapped second-site mutations to NS2 and domain III of NS5A that can compensate the S168G-mediated defects in infectious particle production, and this genetic interaction suggests that NS5A acts at different, temporally distinct steps in the generation of infectious HCV virions.

Finally, a role for NS3 in the assembly of infectious intracellular HCV particles was discovered by the isolation of an adaptive mutation (Q221L) in the NS3 Hel domain. This mutation was required for virion assembly in cells transfected with an intergenotypic chimeric HCV genome containing the core to NS2 region of genotype 1a H77c in the JFH-1 background<sup>247</sup>. The Q221L compensatory mutation did not affect the recruitment



of NS3, core, or NS5A to lipid droplets, and had no effect on the known enzymatic activities of NS3 or viral RNA synthesis<sup>114</sup>. The Q221L substitution enhanced infectious virus production by the JFH1 construct as well, suggesting it might be a cell-culture adaptive mutation and, consequently, interaction of NS3 with a host protein might be important for its function in assembly. However, expression of NS3 containing the Q221L mutation from an alphavirus replicon did not result in *trans* complementation of the virus assembly defects displayed by the original chimeric HCV, implicating NS3 in an early step of virion morphogenesis.

In conclusion, interplay between replicase subunits with catalytic activities important for viral RNA synthesis (such as NS3) and those responsible for inducing membrane rearrangements to accommodate the viral RCs (such as flavivirus NS2A) appears critical for virus production in *Flaviviridae*. Nonstructural proteins seem to be involved in multiple steps of the virion assembly process – from selection of replication- and translation-competent RNA molecules for encapsidation<sup>109,163</sup> to “chaperoning” capsid-RNA interactions<sup>121</sup> and the subsequent association of nucleocapsids with viral glycoproteins, and even in events conferring infectivity to assembled intracellular particles<sup>248</sup>. Functional studies of nonstructural proteins, as well as recent advances in determining the ultrastructure of virus-infected cells, have collectively challenged the view of a spatial and temporal segregation between +RNA virus genome replication and encapsidation. They have also suggested the existence of intimate connections between intracellular compartments for RNA replication and encapsidation, as well as among the protein factors directing both processes. Such a “network” would be exceptionally well suited for ensuring the selective packaging of replication- and translation-competent messenger-sense viral RNA genomes.

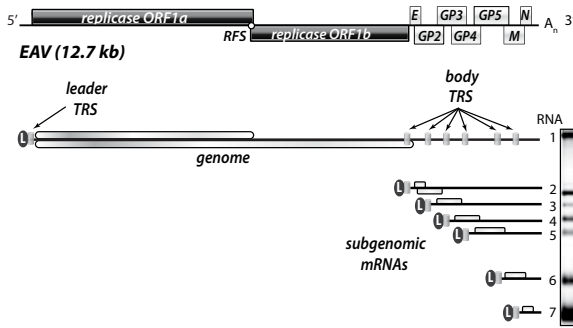
## **WHICH VIRAL AND CELLULAR FACTORS DRIVE THE PROGRESSION OF THE NIDOVIRUS REPLICATIVE CYCLE?**

The *Nidovirales* order comprises the distantly related families *Arteriviridae*, *Coronaviridae* (subfamilies *Coronavirinae* and *Torovirinae* that include three and two genera, respectively), and *Roniviridae* (<http://www.ictvonline.org/virusTaxonomy.asp?version=2009>). The messenger-sense single-stranded nidovirus genomes are polycistronic. Members of *Coronaviridae* and *Roniviridae* have the largest genomes among RNA viruses known to date (25–32 kb), while genome size varies between 12.7 – 15.7 kb in *Arteriviridae*<sup>67</sup>. A multidomain replicase, which contains an array of evolutionarily related core replicative enzyme domains<sup>34,37,68,69,73,183,197</sup>, is encoded in two large 5′-proximal ORFs in nidovirus genomes, while the viral structural genes are located in the 3′-proximal genomic regions. Translation of the 5′-capped, 3′-polyadenylated nidovirus genomic RNA results in two multidomain replicase polyproteins - pp1a and pp1ab, with expression of the latter requiring a -1 ribosomal frameshift. The replicase polyproteins are processed into 13 to 16 nonstructural proteins (nsps) by up to four virus-encoded proteases residing in

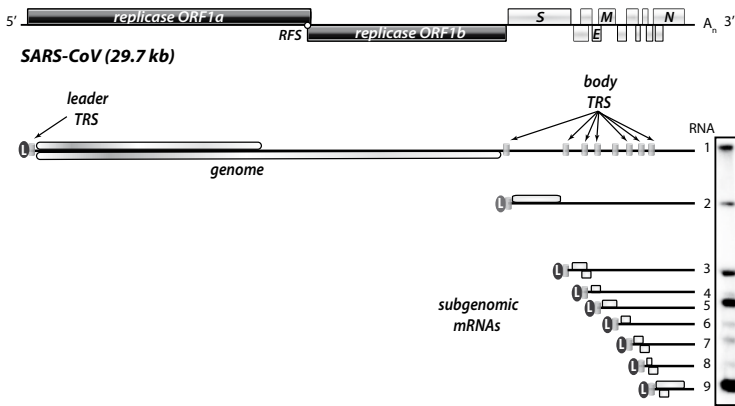




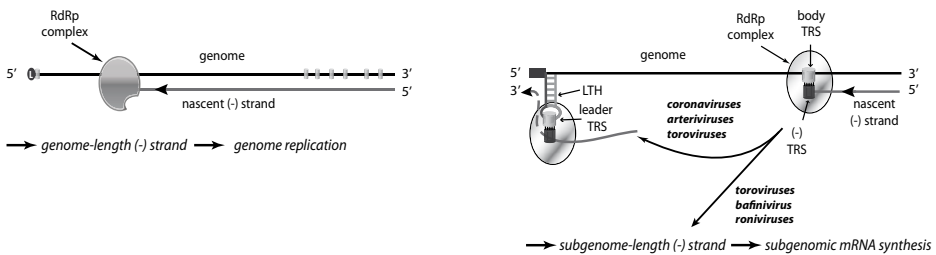
**A**



**B**



**C**



**Figure 4.** Expression of the polycistronic genomes of *Nidovirales* involves the production of an extensive set of sg mRNAs. The genome organization and expression of the arterivirus EAV (A) and the coronavirus SARS-CoV (B) are diagrammed. The 5'-proximal replicase ORFs are designated, as well as the downstream ORFs encoding the viral structural proteins envelope (E), membrane (M), nucleocapsid (N), and glycoproteins (GP) 2-5 for EAV, and E, M, N and spike (S) for SARS-CoV. The 3' poly(A) tail ( $A_n$ ) is also shown. An overview of viral mRNA species produced in infected cells is depicted below each schematic representation of genome organization. The ORFs translated from the respective mRNAs, the 5' leader (L) and the positions of transcription-regulating sequences (TRS) are shown. The gel hybridization images on the right are representative of the wild-type accumulation levels of the seven EAV mRNAs (A) and the nine SARS-CoV mRNAs (B) around the peak of viral RNA synthesis in infected cells. The SARS-CoV gel hybridization image was reproduced from van Hemert *et al.*<sup>223</sup> (C) Model for nidovirus replication and transcription. Continuous minus-strand RNA synthesis yields a genome-length minus strand template for genome replication (left panel). Discontinuous minus-strand RNA synthesis<sup>175</sup> results in a nested

set of subgenome-length minus strands that serve as templates for sg mRNA synthesis in arteri- and coronaviruses. This process is likely carried out by an RdRp complex with a different composition, and is guided by a base pairing interaction between the TRS complement [(-)TRS] at the 3' end of the nascent minus-strand and the genomic leader TRS, present in a RNA hairpin structure (LTH). Subgenomic mRNAs produced by members of the *Roniviridae* family and *Torovirinae* subfamily lack a common leader sequence, with the exception of the largest *Torovirus* sg mRNA<sup>35,226</sup>. The minus-strand templates for sg mRNAs lacking a common leader are thought to be produced via a "premature termination" mechanism similar to that employed by TBSV (see Fig. 2 and main text for details).

pp1a. Nidovirus replicase subunits direct the formation of a membrane-anchored RNA-synthesizing complex that mediates viral genome replication and the synthesis of a 3'-coterminal nested set of subgenomic (sg) mRNAs, the latter being required for expression of the viral structural proteins. Depending on the virus species, a set of 2 to 10 sg mRNAs are produced in nidovirus-infected cells by a process commonly referred to as "transcription" to distinguish it from genome amplification ("replication"). Most of these molecules are structurally polycistronic, but generally serve as templates for translation of only the 5'-proximal ORF, which is not found in the next smaller sg mRNA. Arteri- and coronavirus sg mRNAs have a unique "chimeric" structure, since they contain a common "leader" sequence representing a 5'-proximal fraction of the genome, linked to different "body" segments co-linear with the 3'-proximal part of the genome (Fig. 4, reviewed in<sup>154,176</sup>). By contrast, sg mRNAs produced by members of the *Roniviridae* family and *Torovirus* genus lack a common leader sequence, with the exception of the largest torovirus sg mRNA<sup>35,195,226</sup>.

### **How are active nidovirus RNA-synthesizing complexes assembled and what is their composition?**

Research on the nidovirus replicative cycle has been largely carried out with a number of prototype members that are easily propagated in cell culture, and, more recently, for which full-length cDNA clones have been constructed<sup>4,32,42,217,250,251</sup>. Arguably the best-studied nidovirus prototype, in terms of molecular biology, is the one with the smallest genome (12.7 kb) – the arterivirus prototype equine arteritis virus (EAV), although coronavirus representatives such as murine hepatitis virus (MHV), transmissible gastroenteritis virus (TGEV), and human coronavirus 229E (HCoV 229E) have also been extensively characterized.

Initially, comparative sequence analysis was instrumental in assigning putative functions to the numerous nidovirus nsps. Combined with the isolation of a coronavirus as the etiological agent of severe acute respiratory syndrome (SARS)<sup>43,95</sup>, it has spurred numerous biochemical and structural studies of the core viral enzymes as potential targets for antiviral compounds. Consequently, a wealth of information has been obtained on the properties of individual replicative proteins. In vitro assays for RNA synthesis by recombinant forms of the RdRp domain-containing nsps of arteri- and coronaviruses have recently been developed<sup>17,206</sup>. Along with the availability of enzymatically active, recom-

binant Hel domain-containing nidovirus nsps<sup>81,188</sup>, these assays will constitute important tools for unraveling the fundamental aspects of nidovirus RNA synthesis. Also, apart from the “core” viral enzymes, various nsps were predicted and/or found to contain unusual RNA-processing enzymes that are rare or lacking in other RNA viruses, such as an RNA primase, an exoribonuclease and an endoribonuclease (78,80,127,196 and chapter 4 of this thesis). Advances have also been made in characterizing the architecture of membrane-anchored nidovirus RCs, as well as the roles of membrane-spanning nsps in inducing the extensive membrane alterations characteristic of nidovirus infection<sup>9,31,56,91,148,149,159,167,198</sup>. Our understanding of the assembly and composition of the nidovirus RNA-synthesizing complex, the regulation of its different activities, and the connection of viral RNA synthesis to assembly of infectious particles is, by contrast, only fragmentary.

Nidoviruses employ translational and post-translational control to ensure the temporally and quantitatively regulated expression of the viral replicase subunits, which are produced from two large translation products. The efficiency of the programmed –1 ribosomal frameshifting required for translation of ORF1b<sup>25</sup>, as well as the temporally regulated proteolytic processing of the pp1a and pp1ab precursors, are part of an elaborate mechanism aimed at controlling the relative abundance of nidovirus replicative proteins in infected cells<sup>199,218,220,235,254,255</sup> (for a review see reference<sup>255</sup>). Regulated proteolysis of larger polypeptides, presumably determined by accessibility of cleavage sites and enzyme substrate specificity, may serve to trigger or inhibit specific replicase functions at certain stages of the viral replicative cycle<sup>255</sup>. Although the critical importance of polyprotein processing for nidovirus replication has been well established<sup>90,213,218</sup>, the significance of the numerous (alternative) processing intermediates generated in infected cells for the different stages of the viral replicative cycle is poorly understood.

A classical genetic approach involving detailed analysis of conditional-lethal, temperature-sensitive mutants of MHV has provided valuable insights into the functions of individual viral replicase proteins in viral RNA synthesis. By complementation analysis, this method has also allowed the identification of potential protein–protein interactions that may regulate the activity of MHV RNA-synthesizing complexes<sup>177</sup>. Consequently, ORF1a-encoded replicase subunits were proposed to be *cis*-active, while ORF1b gene products were postulated to function *in trans* following their proteolytic release from pp1ab<sup>177</sup>. Better understanding of the results from such genetic experiments, however, requires rigorous characterization of all steps of the viral replicative cycle. This is exemplified by a temperature-sensitive mutation in MHV nsp10 that implicated this protein in minus-strand RNA production, but was later shown to abrogate polyprotein processing by the main viral proteinase, which is required for viral RNA synthesis<sup>41,90,177</sup>. Numerous pair-wise associations between SARS-CoV proteins have also been identified in large-scale studies using methods such as yeast two-hybrid screens, GST pull-down assays and co-immunoprecipitation<sup>79,150,230</sup>. Whether these interactions indeed occur in the complex intracellular environment of the infected cell, and if so, what their significance is for viral replication, however, remains to be established.



The dynamic associations of viral and/or cellular proteins with regulatory viral RNA sequences, which are likely necessary to ensure the progression of the nidovirus replicative cycle, are slowly emerging. RNA structure probing and reverse genetics studies of *cis*-acting replication signals in the 3' UTRs of arteri- and coronavirus genomes have suggested roles for pseudoknot interactions composed of overlapping RNA elements as regulatory "molecular switches"<sup>18,63,229,241,257</sup>. Interestingly, the severe replication defects associated with a six-nucleotide insertion in a loop of the pseudoknot located in the MHV 3' UTR were suppressed by mutations in nsp8 (the putative viral primase<sup>78</sup>), nsp9 (a presumed RNA-binding protein<sup>47</sup>), and the last nucleotide of the MHV genome upstream of the poly(A) tail<sup>257</sup>. Based on these genetic data, interactions between RNA signals in the 3' UTR of the MHV genome were postulated to form the site of assembly of a complex that includes at least nsp8 and nsp9. Upon substrate recognition, initiation of minus-strand synthesis at the 3' end of the genomic RNA by the nsp8 primase would disrupt the "replicase assembly" RNA signal, thereby allowing formation of the pseudoknot, which may serve to recruit the viral RdRp and Hel and thus trigger elongation of the nascent minus strand. Although appealing, this model does not provide a mechanism for copying sequences from the 3' poly(A) tail, and coronavirus minus-strand RNAs have been shown to contain short poly(U) stretches at their 5' ends<sup>74,174</sup>. The proposed model would also not allow for differentiating between plus strands of genome- or subgenome-length, and the latter presumably do not serve as templates for minus-strand synthesis.

It is worthy of note that circularization of nidovirus genomes has been postulated to facilitate both genome replication and sg mRNA synthesis<sup>103,201,256</sup>. The poly(A) tail was identified as an important *cis*-acting replication element in bovine coronavirus (BCoV), and a specific interaction of PABP with the BCoV 3' UTR, mediated largely by the poly(A) tail sequence, has been described. BCoV DI RNAs with shortened poly(A) tails had virtually identical stabilities in the absence of viral infection, but exhibited weaker PABP binding *in vitro* and also replicated poorly *in vivo*<sup>201</sup>. In analogy to the model proposed for poliovirus<sup>72</sup>, PABP binding to the poly(A) tail of nidovirus genomes might coordinate its amplification, especially in view of the shared dependency of isolated viral replication complexes from poliovirus, EAV, and SARS-CoV on host factors for *in vitro* activity<sup>13,222,223</sup>. So far, experimental support for genome circularization in nidoviruses is scarce, although a recent report provided genetic evidence for an interaction between the leader sequence and the 3' UTR of the MHV genome<sup>103</sup>. In addition, a number of host proteins have been shown to specifically bind to sequences in the plus- and minus-strand RNAs of MHV, though the mechanistic details and significance of these interactions need further investigation<sup>57,64,76,77,191-193,201</sup>.

### **Nidovirus transcription: stop-or-go during minus-strand synthesis?**

Regulation of nidovirus sg mRNA production is one of the most-studied steps of the nidovirus replicative cycle, largely due to the unique mosaic nature of arterivirus and



coronavirus sg mRNAs. According to the now widely supported model proposed by Sawicki and Sawicki<sup>175</sup>, the structure of the arterivirus and coronavirus sg mRNAs derives from a discontinuous step during minus-strand RNA synthesis, which is guided by conserved, AU-rich transcription-regulating sequences (TRS)<sup>156,224,256</sup>. “Body TRS” motifs are found preceding almost all structural protein ORFs, and a “leader TRS” is present at the 3′ end of the genomic leader sequence. Minus-strand RNA synthesis, initiated at the 3′ end of the viral genome, is presumably attenuated at one of the body TRS regions and the nascent minus strand, carrying the body TRS complement at its 3′-end, is translocated to the 5′-proximal region of the genomic template. This step is guided by base-pairing between the genomic leader TRS and the copy of the body TRS present in the 3′ end of the nascent minus strand. Nascent strands are subsequently extended with the complement of the genomic leader sequence, yielding a nested set of subgenome-length minus-strand templates that can be directly copied into the various sg mRNAs (reviewed in references<sup>154,176</sup>). Alternatively, the successive body TRS motifs can all be ignored during minus-strand RNA synthesis, resulting in a full-length complement of the genome which serves as a template for genome amplification.

The switch between “continuous” and “discontinuous” minus-strand RNA synthesis is therefore crucial for the production of viral structural proteins. Moreover, the amounts of structural proteins are presumably determined by the abundance of their mRNA templates, since the diverse viral mRNA species are produced in nonequimolar amounts, and their relative molar ratios are largely constant both in coronavirus- and arterivirus-infected cells<sup>12,38,174</sup>. In the model of discontinuous extension during minus-strand synthesis, the relative abundance of a sg mRNA would be determined by the degree of “attenuation” of RNA synthesis at the corresponding body TRS, and consequently, by the abundance of its minus-strand template.

Extensive research on the TRSs in arterivirus and coronavirus genomes has yielded important insights into the primary and higher-order structure determinants of sg mRNA production. The base-pairing potential between conserved “core” motifs in the leader TRS and the body TRS complement in the nascent minus strand is a key determinant of sg mRNA production, but also the sequence context of body TRS motifs and their proximity to the genomic 3′ end, reflected in a “gradient” of sg RNA sizes, can influence the accumulation of viral mRNA species (<sup>155-157,200,224,256</sup>, reviewed in references<sup>154,176</sup>). Moreover, a leader TRS hairpin (LTH), containing the leader TRS in its loop, is essential for EAV sg mRNA synthesis<sup>214,214,215</sup>. The LTH likely serves as a transcription-guiding element, since it could still direct the production of leader-containing sg mRNAs even when placed at another position in the 5′-proximal region of the EAV genome<sup>215</sup>. Although important, sequence identity between the leader TRS and body TRS can be insufficient for directing sg mRNA production<sup>156</sup>.

The role of higher-order RNA structures in attenuation of minus-strand synthesis at body TRS regions remains unclear. Recent evidence suggests a long-range RNA-RNA interaction enhances production of the N-coding sg mRNA of TGEV<sup>135</sup>. A mode of regulation of minus-strand attenuation by a network of RNA signals, similar to the one described for

TBSV<sup>104</sup>, might therefore have evolved in arteriviruses and coronaviruses, although it is probably even more sophisticated, since it likely includes the LTH and needs to direct the balanced synthesis of a larger set of sg transcripts. Physical association between the 5' and 3' ends of the coronavirus genome has been proposed to position nascent minus strands in the proximity of the leader sequence, facilitating base-pairing between the leader TRS and the body TRS complement and thus strand transfer during minus-strand synthesis<sup>256</sup>. Apart from the recently described genetic interaction between the 5' and 3' UTRs of MHV<sup>103</sup>, however, there is currently no other evidence in support of genome circularization and its potential importance for sg mRNA synthesis in arteriviruses and coronaviruses.

The protein factors governing the production of subgenome-length minus strands by interactions with the transcription-regulatory RNA networks are, with a few prominent exceptions, largely unexplored. A serendipitous discovery of a mutation in nsp10, the Hel domain-containing replicase subunit, revealed that sg mRNA synthesis can be functionally uncoupled from genome replication in EAV<sup>217</sup>. The S2429P substitution (amino acid position in EAV pp1ab) did not interfere with replicase polyprotein processing or genomic RNA accumulation, but resulted in a more than 100-fold reduction of subgenome-length minus-strand and plus-strand levels<sup>225</sup>. S2429 is located just downstream of a predicted zinc-binding subdomain in nsp10 and upstream of the conserved Hel motif, and the S2429P mutation had no detectable effect on the *in vitro* helicase activity of recombinant EAV nsp10. Moreover, despite extensive mutagenesis of EAV nsp10, substitutions of other residues resulting in similar specific defects in sg mRNA production have not been described, while deletion of the entire nsp10-coding sequence or mutations in the Hel active site abolished all viral RNA synthesis<sup>187,219</sup>.

By contrast, deleting the coding sequence of another replicase subunit, nsp1, from the EAV genome specifically reduced the accumulation of subgenome-length viral RNAs of both polarities to levels undetectable by RT-PCR, while genome replication was largely unaffected<sup>212</sup>. Subgenomic mRNA production could be partially restored by IRES-mediated expression of nsp1 from an alternative region in the EAV genome or by plasmid-driven nsp1 expression (<sup>212</sup>, D.D. Nedialkova, unpublished observations), demonstrating the protein is able to function *in trans*. Substitutions of amino acids in the zinc finger subdomain, but also in the two papain-like cysteine proteinase subdomains of nsp1 selectively abolished sg mRNA accumulation (Chapters 5 and 6 of this thesis). Evidence for a specific role in sg mRNA production was also obtained for an arterivirus ortholog of EAV nsp1 - nsp1 $\alpha$  of porcine reproductive and respiratory syndrome virus (PRRSV)<sup>94</sup>.

Most notably, EAV nsp1 was found to modulate the accumulation of viral minus-strand templates to control the relative abundance of both genome-length and subgenome-length mRNAs in EAV-infected cells. Balanced EAV mRNA accumulation, in turn, was demonstrated to be critical for efficient production of infectious progeny particles (Chapters 5 and 6). These data provided evidence not only for transcriptional regulation as a genuine mechanism for controlling the relative abundance of viral proteins, but for



the importance of this control for the later stages of EAV infection, i.e. virion assembly. Interestingly, certain mutations in nsp1 were also found to affect virus production, but without significantly disturbing viral RNA synthesis (Chapters 5 and 6 of this thesis). Also, although nsp1 itself is dispensable for EAV genome replication, blocking the autoproteolytic release of the protein from the nascent replicase polypeptide reduced the synthesis of all viral RNAs to undetectable levels (chapter 5). A single arterivirus protein thus plays a decisive role in the integral control of replication, sg mRNA synthesis, and virus production, and studies on EAV nsp1 have provided the most substantial evidence to date for coupling of these processes in the nidovirus replicative cycle.

A specific role in controlling sg mRNA production has not been assigned to any coronavirus protein so far, and uncoupling of “continuous” from “discontinuous” coronavirus RNA synthesis has only been achieved by mutating TRS motifs<sup>256</sup>. Unlike the EAV reverse genetics system, however, currently available reverse genetics systems for coronaviruses suffer from low transfection efficiencies, likely stemming from difficulties to obtain high-quality *in vitro*-synthesized coronavirus full-length RNAs due to their large sizes. Coronavirus phenotypes are thus frequently analyzed after passaging of the mutant viruses, an experimental setup that precludes the identification and characterization of replication-competent, transcription-negative mutants.

### **How is nidovirus particle assembly orchestrated?**

Mechanisms of virion biogenesis employed by the different nidovirus groups probably differ, due to their dissimilar virion morphologies and the widely different sets of structural proteins they encode. For example, arteriviruses capsids are probably icosahedral, while those of coronaviruses have a helical structure<sup>11,142</sup>, and the membrane-spanning structural proteins of nidoviruses are also very divergent. Nevertheless, the nucleocapsid (N) proteins of both EAV and MHV partially co-localize with viral replicase subunits at the presumed sites of viral RNA synthesis, especially in the early stages of infection, suggesting a topological connection between the sites of genome replication and packaging<sup>211,216</sup>.

RNA motifs in the genomes of MHV, BCoV and TGEV have been proposed to mediate the specific encapsidation of genomic RNAs. Insertion of those “packaging signals” into nonviral transcripts led to their specific incorporation into viral particles, although with variable efficiency<sup>33,49,119,242</sup>. A selective interaction between the membrane (M) protein of MHV and RNA molecules containing the packaging signal was proposed to be critical for the specificity and selectivity of RNA encapsidation<sup>140</sup> and interactions between the M protein and the envelope (E) protein of coronaviruses were implicated in determining the intracellular sites of virus assembly and budding<sup>228</sup>.

Finally, although a number of studies have examined the sequence requirements for packaging of nidovirus DI RNAs<sup>33,36,49,82,129</sup>, a link between genome translation, replication, and encapsidation, similar to those described for PV and KUN virus<sup>88,146</sup> has not been systematically investigated for any nidovirus. Likewise, the significance of nidovirus

replicase subunits for infectious particle production remains largely unexplored, and to date a role in virion assembly has been demonstrated only for EAV nsp1 (Chapters 5 and 6 of this thesis). It is interesting to note that the N protein was shown to facilitate coronavirus genome replication by an as yet uncharacterized mechanism<sup>3,180,209</sup>. This unusual role of a structural protein in RNA synthesis might not be a general feature of nidovirus replication, since all structural proteins were shown to be dispensable for EAV genome replication and sg mRNA synthesis<sup>130</sup>.

### **Taking nidovirus replication apart: lessons learned from other +RNA viruses**

In the past two decades, research on the spatio-temporal regulation of the eukaryotic +RNA virus replicative cycle has led to the emergence of several regulatory strategies commonly employed by virus families with diverse genetic plans and gene expression modes. A considerable degree of integration between translation, replication, and encapsidation of the genomic RNA seems to have evolved in diverse +RNA virus groups<sup>88,144,146</sup>. Plus-strand RNA viruses from different families seem to utilize remarkably analogous “molecular switch” mechanisms to regulate the use of the genomic RNA as a template for translation or minus-strand RNA synthesis. Circularization of the viral genome has been implicated in ensuring the switch between these two processes in virus families as diverse as *Picornaviridae*, *Flaviviridae* and *Tombusviridae*<sup>51,53,72,170</sup>. The importance of long-distance RNA-RNA interactions, often spanning thousands of nucleotides in +RNA viral genomes, for the formation of a variety of functional regulatory RNA elements is slowly emerging<sup>40,104,184,243</sup>. These findings have major consequences for research on regulatory RNA motifs, which are traditionally viewed as “localized” sequences or structures. In addition, studies on +RNA virus replication in genetically tractable model hosts have provided ample evidence for essential roles of cellular proteins in many of the fundamental steps of the viral replicative cycle<sup>2,10,27,72,97,120,165,186,232</sup>. Finally, recent data have revealed that the intracellular sites of viral RNA synthesis and virion assembly are interconnected, and have also implicated common viral protein factors into governing these processes<sup>137,238</sup>.

The emergent common themes in +RNA virus replication should undoubtedly be reckoned with in future studies on regulatory events in the nidovirus replicative cycle. For example, nidovirus RdRp attenuation during minus-strand RNA synthesis at body TRS regions may mechanistically resemble the “premature termination” step in tombusviruses and may also be controlled by a network of long-distance RNA-RNA interactions<sup>135</sup>. Long-range RNA-RNA interactions playing a role in the generation of subgenome-length minus strands during nidovirus infection may be the reason that no conserved “local” higher-order structures have been identified in the body TRS regions of nidovirus genomes so far. This possibility should be accounted for when designing *in vitro* assays to examine the possible stalling of nidovirus RdRps at body TRS motifs, as well as in experiments aimed at identifying protein factors interacting with TRS networks. Recent reports on the characterization of arteri- and coronavirus RNA-synthesizing complexes isolated from infected cells have shown they require cytoplasmic host factor(s) for activity *in vitro*<sup>222,223</sup>,



confirming a key role of host proteins in nidovirus replication. The identity of these factors and the mechanisms of their action remain to be identified, but they may be part of a larger pool of cellular proteins recruited to facilitate nidovirus replication. Pinpointing the step(s) of nidovirus RNA synthesis regulated by cellular factors will require an expansion of the current set of experimental approaches with assays that allow the analysis of viral RNA translation and minus-strand RNA synthesis at a level of detail that is currently not attainable. Lastly, research on the determinants of nidovirus particle assembly may benefit from exploring a potential connection between nidovirus genome replication and encapsidation. Such a link could ensure the efficiency of RNA incorporation into progeny particles, for which the presence of a packaging signal is seemingly insufficient<sup>24</sup>, and may shed more light on the possible roles of nidovirus replicative proteins in virion biogenesis.

## REFERENCES

1. **Agapov, E. V., C. L. Murray, I. Frolov, L. Qu, T. M. Myers, and C. M. Rice. 2004.** Uncleaved NS2-3 is required for production of infectious bovine viral diarrhea virus. *J.Virol.* **78**:2414-2425
2. **Ahlquist, P., A. O. Noueiry, W. M. Lee, D. B. Kushner, and B. T. Dye. 2003.** Host factors in positive-strand RNA virus genome replication. *J.Virol.* **77**:8181-8186
3. **Almazan, F., C. Galan, and L. Enjuanes. 2004.** The nucleoprotein is required for efficient coronavirus genome replication. *J.Virol.* **78**:12683-12688
4. **Almazan, F., J. M. Gonzalez, Z. Penzes, A. Izeta, E. Calvo, J. Plana-Duran, and L. Enjuanes. 2000.** Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. *Proc.Natl.Acad.Sci.U.S.A* **97**:5516-5521
5. **Andino, R., G. E. Rieckhof, P. L. Achacoso, and D. Baltimore. 1993.** Poliovirus RNA synthesis utilizes an RNP complex formed around the 5'-end of viral RNA. *EMBO J.* **12**:3587-3598
6. **Andino, R., G. E. Rieckhof, and D. Baltimore. 1990.** A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. *Cell* **63**:369-380
7. **Andino, R., G. E. Rieckhof, D. Trono, and D. Baltimore. 1990.** Substitutions in the protease (3Cpro) gene of poliovirus can suppress a mutation in the 5' noncoding region. *J.Virol.* **64**:607-612
8. **Appel, N., M. Zayas, S. Miller, J. Krijnse-Locker, T. Schaller, P. Friebe, S. Kallis, U. Engel, and R. Bartenschlager. 2008.** Essential role of domain III of nonstructural protein 5A for hepatitis C virus infectious particle assembly. *PLoS.Pathog.* **4**:e1000035
9. **Baliji, S., S. A. Cammer, B. Sobral, and S. C. Baker. 2009.** Detection of nonstructural protein 6 in murine coronavirus-infected cells and analysis of the transmembrane topology by using bioinformatics and molecular approaches. *J.Virol.* **83**:6957-6962
10. **Barajas, D., Y. Jiang, and P. D. Nagy. 2009.** A unique role for the host ESCRT proteins in replication of Tomato bushy stunt virus. *PLoS.Pathog.* **5**:e1000705
11. **Barcena, M., G. T. Oostergetel, W. Bartelink, F. G. Faas, A. Verkleij, P. J. Rottier, A. J. Koster, and B. J. Bosch. 2009.** Cryo-electron tomography of mouse hepatitis virus: Insights into the structure of the coronavirus. *Proc.Natl.Acad.Sci.U.S.A* **106**:582-587
12. **Baric, R. S. and B. Yount. 2000.** Subgenomic negative-strand RNA function during mouse hepatitis virus infection. *J.Virol.* **74**:4039-4046
13. **Barton, D. J., E. P. Black, and J. B. Flanagan. 1995.** Complete replication of poliovirus in vitro: preinitiation RNA replication complexes require soluble cellular factors for the synthesis of VPg-linked RNA. *J.Virol.* **69**:5516-5527
14. **Barton, D. J. and J. B. Flanagan. 1993.** Coupled translation and replication of poliovirus RNA in vitro: synthesis of functional 3D polymerase and infectious virus. *J.Virol.* **67**:822-831
15. **Barton, D. J., B. J. Morasco, and J. B. Flanagan. 1999.** Translating ribosomes inhibit poliovirus negative-strand RNA synthesis. *J.Virol.* **73**:10104-10112
16. **Bedard, K. M., B. L. Walter, and B. L. Semler. 2004.** Multimerization of poly(rC) binding protein 2 is required for translation initiation mediated by a viral IRES. *RNA.* **10**:1266-1276
17. **Beerens, N., B. Selisko, S. Ricagno, I. Imbert, L. van der Zanden, E. J. Snijder, and B. Canard. 2007.** De novo initiation of RNA synthesis by the arterivirus RNA-dependent RNA polymerase. *J.Virol.* **81**:8384-8395
18. **Beerens, N. and E. J. Snijder. 2007.** An RNA pseudoknot in the 3' end of the arterivirus genome has a critical role in regulating viral RNA synthesis. *J.Virol.* **81**:9426-9436
19. **Bienz, K., D. Egger, T. Pfister, and M. Troxler. 1992.** Structural and functional characterization of the poliovirus replication complex. *J.Virol.* **66**:2740-2747

20. **Blight, K. J., J. A. McKeating, and C. M. Rice. 2002.** Highly permissive cell lines for subgenomic and genomic hepatitis C virus RNA replication. *J.Virol.* **76**:13001-13014
21. **Blyn, L. B., K. M. Swiderek, O. Richards, D. C. Stahl, B. L. Semler, and E. Ehrenfeld. 1996.** Poly(rC) binding protein 2 binds to stem-loop IV of the poliovirus RNA 5' noncoding region: identification by automated liquid chromatography-tandem mass spectrometry. *Proc.Natl.Acad. Sci.U.S.A* **93**:11115-11120
22. **Blyn, L. B., J. S. Towner, B. L. Semler, and E. Ehrenfeld. 1997.** Requirement of poly(rC) binding protein 2 for translation of poliovirus RNA. *J.Virol.* **71**:6243-6246
23. **Borman, A. M., F. G. Deliat, and K. M. Kean. 1994.** Sequences within the poliovirus internal ribosome entry segment control viral RNA synthesis. *EMBO J.* **13**:3149-3157
24. **Bos, E. C., J. C. Dobbe, W. Luytjes, and W. J. Spaan. 1997.** A subgenomic mRNA transcript of the coronavirus mouse hepatitis virus strain A59 defective interfering (DI) RNA is packaged when it contains the DI packaging signal. *J.Virol.* **71**:5684-5687
25. **Brierley, I., P. Digard, and S. C. Inglis. 1989.** Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* **57**:537-547
26. **Brohm, C., E. Steinmann, M. Friesland, I. C. Lorenz, A. Patel, F. Penin, R. Bartenschlager, and T. Pietschmann. 2009.** Characterization of determinants important for hepatitis C virus p7 function in morphogenesis by using trans-complementation. *J.Virol.* **83**:11682-11693
27. **Castorena, K. M., S. A. Weeks, K. A. Stapleford, A. M. Cadwallader, and D. J. Miller. 2007.** A functional heat shock protein 90 chaperone is essential for efficient flock house virus RNA polymerase synthesis in *Drosophila* cells. *J.Virol.* **81**:8412-8420
28. **Choi, I. R., M. Ostrovsky, G. Zhang, and K. A. White. 2001.** Regulatory activity of distal and core RNA elements in Tombusvirus subgenomic mRNA2 transcription. *J.Biol.Chem.* **276**:41761-41768
29. **Choi, I. R. and K. A. White. 2002.** An RNA activator of subgenomic mRNA1 transcription in tomato bushy stunt virus. *J.Biol.Chem.* **277**:3760-3766
30. **Clarke, D., S. Griffin, L. Beales, C. S. Gelais, S. Burgess, M. Harris, and D. Rowlands. 2006.** Evidence for the formation of a heptameric ion channel complex by the hepatitis C virus p7 protein in vitro. *J.Biol.Chem.* **281**:37057-37068
31. **Clementz, M. A., A. Kanjanahaluethai, T. E. O'Brien, and S. C. Baker. 2008.** Mutation in murine coronavirus replication protein nsp4 alters assembly of double membrane vesicles. *Virology* **375**:118-129
32. **Coley, S. E., E. Lavi, S. G. Sawicki, L. Fu, B. Schelle, N. Karl, S. G. Siddell, and V. Thiel. 2005.** Recombinant mouse hepatitis virus strain A59 from cloned, full-length cDNA replicates to high titers in vitro and is fully pathogenic in vivo. *J.Virol.* **79**:3097-3106
33. **Cologna, R., J. F. Spagnolo, and B. G. Hogue. 2000.** Identification of nucleocapsid binding sites within coronavirus-defective genomes. *Virology* **277**:235-249
34. **Cowley, J. A., C. M. Dimmock, K. M. Spann, and P. J. Walker. 2000.** Gill-associated virus of *Penaeus monodon* prawns: an invertebrate virus with ORF1a and ORF1b genes related to arteri- and coronaviruses. *J.Gen.Virol.* **81**:1473-1484
35. **Cowley, J. A., C. M. Dimmock, and P. J. Walker. 2002.** Gill-associated nidovirus of *Penaeus monodon* prawns transcribes 3'-coterminal subgenomic mRNAs that do not possess 5'-leader sequences. *J.Gen.Virol.* **83**:927-935
36. **Dalton, K., R. Casais, K. Shaw, K. Stirrups, S. Evans, P. Britton, T. D. Brown, and D. Cavanagh. 2001.** cis-acting sequences required for coronavirus infectious bronchitis virus defective-RNA replication and packaging. *J.Virol.* **75**:125-133



37. **den Boon, J. A., E. J. Snijder, E. D. Chirnside, A. A. de Vries, M. C. Horzinek, and W. J. Spaan. 1991.** Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. *J.Virol.* **65**:2910-2920
38. **den Boon, J. A., W. J. M. Spaan, and E. J. Snijder. 1995.** Equine arteritis virus subgenomic RNA transcription - UV inactivation and translation inhibition studies. *Virology* **213**:364-372
39. **Dentzer, T. G., I. C. Lorenz, M. J. Evans, and C. M. Rice. 2009.** Determinants of the hepatitis C virus nonstructural protein 2 protease domain required for production of infectious virus. *J.Virol.* **83**:12702-12713
40. **Diviney, S., A. Tuplin, M. Struthers, V. Armstrong, R. M. Elliott, P. Simmonds, and D. J. Evans. 2008.** A hepatitis C virus cis-acting replication element forms a long-range RNA-RNA interaction with upstream RNA sequences in NS5B. *J.Virol.* **82**:9008-9022
41. **Donaldson, E. F., A. C. Sims, R. L. Graham, M. R. Denison, and R. S. Baric. 2007.** Murine hepatitis virus replicase protein nsp10 is a critical regulator of viral RNA synthesis. *J.Virol.* **81**:6356-6368
42. **Donaldson, E. F., B. Yount, A. C. Sims, S. Burkett, R. J. Pickles, and R. S. Baric. 2008.** Systematic assembly of a full-length infectious clone of human coronavirus NL63. *J.Virol.* **82**:11948-11957
43. **Drosten, C., S. Gunther, W. Preiser, S. van der Werf, H. R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R. A. Fouchier, A. Berger, A. M. Burguiere, J. Cinatl, M. Eickmann, N. Escriou, K. Grywna, S. Kramme, J. C. Manuguerra, S. Muller, V. Rickerts, M. Sturmer, S. Vieth, H. D. Klenk, A. D. Osterhaus, H. Schmitz, and H. W. Doerr. 2003.** Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N.Engl.J.Med.* **348**:1967-1976
44. **Egger, D., N. Teterina, E. Ehrenfeld, and K. Bienz. 2000.** Formation of the poliovirus replication complex requires coupled viral translation, vesicle production, and viral RNA synthesis. *J.Virol.* **74**:6570-6580
45. **Egger, D., B. Wolk, R. Gosert, L. Bianchi, H. E. Blum, D. Moradpour, and K. Bienz. 2002.** Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. *J.Virol.* **76**:5974-5984
46. **Egloff, M. P., D. Benarroch, B. Selisko, J. L. Romette, and B. Canard. 2002.** An RNA cap (nucleoside-2'-O-)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and functional characterization. *EMBO J.* **21**:2757-2768
47. **Egloff, M. P., F. Ferron, V. Campanacci, S. Longhi, C. Rancurel, H. Dutartre, E. J. Snijder, A. E. Gorbalenya, C. Cambillau, and B. Canard. 2004.** The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. *Proc.Natl.Acad.Sci.U.S.A* **101**:3792-3796
48. **Erbel, P., N. Schiering, A. D'Arcy, M. Renatus, M. Kroemer, S. P. Lim, Z. Yin, T. H. Keller, S. G. Vasudevan, and U. Hommel. 2006.** Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nat.Struct.Mol.Biol.* **13**:372-373
49. **Escors, D., A. Izeta, C. Capiscol, and L. Enjuanes. 2003.** Transmissible gastroenteritis coronavirus packaging signal is located at the 5' end of the virus genome. *J.Virol.* **77**:7890-7902
50. **Fabian, M. R. and K. A. White. 2004.** 5'-3' RNA-RNA interaction facilitates cap- and poly(A) tail-independent translation of tomato bushy stunt virus mRNA: a potential common mechanism for tombusviridae. *J.Biol.Chem.* **279**:28862-28872
51. **Fabian, M. R. and K. A. White. 2006.** Analysis of a 3'-translation enhancer in a tombusvirus: a dynamic model for RNA-RNA interactions of mRNA termini. *RNA* **12**:1304-1314
52. **Falgout, B., M. Pethel, Y. M. Zhang, and C. J. Lai. 1991.** Both nonstructural proteins NS2B and NS3 are required for the proteolytic processing of dengue virus nonstructural proteins. *J.Virol.* **65**:2467-2475

53. **Filomatori, C. V., M. F. Lodeiro, D. E. Alvarez, M. M. Samsa, L. Pietrasanta, and A. V. Gamarnik. 2006.** A 5' RNA element promotes dengue virus RNA synthesis on a circular genome. *Genes Dev.* **20**:2238-2249
54. **Flanegan, J. B., R. F. Petterson, V. Ambros, N. J. Hewlett, and D. Baltimore. 1977.** Covalent linkage of a protein to a defined nucleotide sequence at the 5'-terminus of virion and replicative intermediate RNAs of poliovirus. *Proc.Natl.Acad.Sci.U.S.A* **74**:961-965
55. **Franck, N., S. J. Le, C. Guguen-Guillouzo, and L. Erdtmann. 2005.** Hepatitis C virus NS2 protein is phosphorylated by the protein kinase CK2 and targeted for degradation to the proteasome. *J.Virol.* **79**:2700-2708
56. **Gadlage, M. J., J. S. Sparks, D. C. Beachboard, R. G. Cox, J. D. Doyle, C. C. Stobart, and M. R. Denison. 2010.** Murine hepatitis virus nonstructural protein 4 regulates virus-induced membrane modifications and replication complex function. *J.Virol.* **84**:280-290
57. **Galan, C., I. Sola, A. Nogales, B. Thomas, A. Akoulitchev, L. Enjuanes, and F. Almazan. 2009.** Host cell proteins interacting with the 3' end of TGEV coronavirus genome influence virus replication. *Virology* **391**:304-314
58. **Gamarnik, A. V. and R. Andino. 1996.** Replication of poliovirus in *Xenopus* oocytes requires two human factors. *EMBO J.* **15**:5988-5998
59. **Gamarnik, A. V. and R. Andino. 1997.** Two functional complexes formed by KH domain containing proteins with the 5' noncoding region of poliovirus RNA. *RNA* **3**:882-892
60. **Gamarnik, A. V. and R. Andino. 1998.** Switch from translation to RNA replication in a positive-stranded RNA virus. *Genes Dev.* **12**:2293-2304
61. **Gamarnik, A. V. and R. Andino. 2000.** Interactions of viral protein 3CD and poly(rC) binding protein with the 5' untranslated region of the poliovirus genome. *J.Virol.* **74**:2219-2226
62. **Gastaminza, P., S. B. Kapadia, and F. V. Chisari. 2006.** Differential biophysical properties of infectious intracellular and secreted hepatitis C virus particles. *J.Virol.* **80**:11074-11081
63. **Goebel, S. J., B. Hsue, T. F. Dombrowski, and P. S. Masters. 2004.** Characterization of the RNA components of a putative molecular switch in the 3' untranslated region of the murine coronavirus genome. *J.Virol.* **78**:669-682
64. **Goebel, S. J., T. B. Miller, C. J. Bennett, K. A. Bernard, and P. S. Masters. 2007.** A hypervariable region within the 3' cis-acting element of the murine coronavirus genome is nonessential for RNA synthesis but affects pathogenesis. *J.Virol.* **81**:1274-1287
65. **Goodfellow, I. G., D. Kerrigan, and D. J. Evans. 2003.** Structure and function analysis of the poliovirus cis-acting replication element (CRE). *RNA.* **9**:124-137
66. **Goodfellow, I. G., C. Polacek, R. Andino, and D. J. Evans. 2003.** The poliovirus 2C cis-acting replication element-mediated uridylylation of VPg is not required for synthesis of negative-sense genomes. *J.Gen.Virol.* **84**:2359-2363
67. **Gorbalenya, A. E., L. Enjuanes, J. Ziebuhr, and E. J. Snijder. 2006.** Nidovirales: evolving the largest RNA virus genome. *Virus Res.* **117**:17-37
68. **Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov. 1988.** A novel superfamily of nucleoside triphosphate-binding motif containing proteins which are probably involved in duplex unwinding in DNA and RNA replication and recombination. *FEBS Lett.* **235**:16-24
69. **Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov. 1989.** Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Res.* **17**:4847-4861
70. **Gosert, R., D. Egger, V. Lohmann, R. Bartenschlager, H. E. Blum, K. Bienz, and D. Moradpour. 2003.** Identification of the hepatitis C virus RNA replication complex in Huh-7 cells harboring subgenomic replicons. *J.Virol.* **77**:5487-5492



71. **Griffin, S. D., L. P. Beales, D. S. Clarke, O. Worsfold, S. D. Evans, J. Jaeger, M. P. Harris, and D. J. Rowlands. 2003.** The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. *FEBS Lett.* **535**:34-38
72. **Herold, J. and R. Andino. 2001.** Poliovirus RNA replication requires genome circularization through a protein-protein bridge. *Mol.Cell* **7**:581-591
73. **Hodgman, T. C. 1988.** A new superfamily of replicative proteins. *Nature* **333**:22-23
74. **Hofmann, M. A. and D. A. Brian. 1991.** The 5' end of coronavirus minus-strand RNAs contains a short poly(U) tract. *J.Virol.* **65**:6331-6333
75. **Huang, L., J. Hwang, S. D. Sharma, M. R. Hargittai, Y. Chen, J. J. Arnold, K. D. Raney, and C. E. Cameron. 2005.** Hepatitis C virus nonstructural protein 5A (NS5A) is an RNA-binding protein. *J.Biol.Chem.* **280**:36417-36428
76. **Huang, P. and M. M. Lai. 1999.** Polypyrimidine tract-binding protein binds to the complementary strand of the mouse hepatitis virus 3' untranslated region, thereby altering RNA conformation. *J.Virol.* **73**:9110-9116
77. **Huang, P. and M. M. Lai. 2001.** Heterogeneous nuclear ribonucleoprotein a1 binds to the 3'-untranslated region and mediates potential 5'-3'-end cross talks of mouse hepatitis virus RNA. *J.Virol.* **75**:5009-5017
78. **Imbert, I., J. C. Guillemot, J. M. Bourhis, C. Bussetta, B. Coutard, M. P. Egloff, F. Ferron, A. E. Gorbalenya, and B. Canard. 2006.** A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J.* **25**:4933-4942
79. **Imbert, I., E. J. Snijder, M. Dimitrova, J. C. Guillemot, P. Lecine, and B. Canard. 2008.** The SARS-Coronavirus PLnc domain of nsp3 as a replication/transcription scaffolding protein. *Virus Res.* **133**:136-148
80. **Ivanov, K. A., T. Hertzog, M. Rozanov, S. Bayer, V. Thiel, A. E. Gorbalenya, and J. Ziebuhr. 2004.** Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc.Natl.Acad. Sci.U.S.A* **101**:12694-12699
81. **Ivanov, K. A., V. Thiel, J. C. Dobbe, Y. van der Meer, E. J. Snijder, and J. Ziebuhr. 2004.** Multiple enzymatic activities associated with Severe acute respiratory syndrome coronavirus helicase. *J. Virol.* **78**:5619-5632
82. **Izeta, A., C. Smerdou, S. Alonso, Z. Penzes, A. Mendez, J. Plana-Duran, and L. Enjuanes. 1999.** Replication and packaging of transmissible gastroenteritis coronavirus-derived synthetic minigenomes. *J.Virol.* **73**:1535-1545
83. **Jiang, Y., E. Serviene, J. Gal, T. Panavas, and P. D. Nagy. 2006.** Identification of essential host factors affecting tombusvirus RNA replication based on the yeast Tet promoters Hughes Collection. *J.Virol.* **80**:7394-7404
84. **Jirasko, V., R. Montserret, N. Appel, A. Janvier, L. Eustachi, C. Brohm, E. Steinmann, T. Pietschmann, F. Penin, and R. Bartenschlager. 2008.** Structural and functional characterization of nonstructural protein 2 for its role in hepatitis C virus assembly. *J.Biol.Chem.* **283**:28546-28562
85. **Jones, C. T., C. L. Murray, D. K. Eastman, J. Tassello, and C. M. Rice. 2007.** Hepatitis C virus p7 and NS2 proteins are essential for production of infectious virus. *J.Virol.* **81**:8374-8383
86. **Jones, C. T., C. G. Patkar, and R. J. Kuhn. 2005.** Construction and applications of yellow fever virus replicons. *Virology* **331**:247-259
87. **Kempf, B. J. and D. J. Barton. 2008.** Poly(rC) binding proteins and the 5' cloverleaf of uncapped poliovirus mRNA function during de novo assembly of polysomes. *J.Virol.* **82**:5835-5846
88. **Khromykh, A. A., A. N. Varnavski, P. L. Sedlak, and E. G. Westaway. 2001.** Coupling between replication and packaging of flavivirus RNA: evidence derived from the use of DNA-based full-length cDNA clones of Kunjin virus. *J.Virol.* **75**:4633-4640

89. **Khromykh, A. A., A. N. Varnavski, and E. G. Westaway. 1998.** Encapsidation of the flavivirus kunjin replicon RNA by using a complementation system providing Kunjin virus structural proteins in trans. *J.Virol.* **72**:5967-5977
90. **Kim, J. C., R. A. Spence, P. F. Currier, X. Lu, and M. R. Denison. 1995.** Coronavirus protein processing and RNA synthesis is inhibited by the cysteine proteinase inhibitor E64d. *Virology* **208**:1-8
91. **Knoops, K., M. Kikkert, S. H. Worm, J. C. Zevenhoven-Dobbe, Y. van der Meer, A. J. Koster, A. M. Mommaas, and E. J. Snijder. 2008.** SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS.Biol.* **6**:e226
92. **Krausslich, H. G., M. J. Nicklin, C. K. Lee, and E. Wimmer. 1988.** Polyprotein processing in picornavirus replication. *Biochimie* **70**:119-130
93. **Krieger, N., V. Lohmann, and R. Bartenschlager. 2001.** Enhancement of hepatitis C virus RNA replication by cell culture-adaptive mutations. *J.Virol.* **75**:4614-4624
94. **Kroese, M. V., J. C. Zevenhoven-Dobbe, J. N. Bos-de Ruijter, B. P. Peeters, J. J. Meulenberg, L. A. Cornelissen, and E. J. Snijder. 2008.** The nsp1alpha and nsp1 papain-like autoproteases are essential for porcine reproductive and respiratory syndrome virus RNA synthesis. *J.Gen.Virol.* **89**:494-499
95. **Ksiazek, T. G., D. Erdman, C. S. Goldsmith, S. R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J. A. Comer, W. Lim, P. E. Rollin, S. F. Dowell, A. E. Ling, C. D. Humphrey, W. J. Shieh, J. Guarner, C. D. Paddock, P. Rota, B. Fields, J. DeRisi, J. Y. Yang, N. Cox, J. M. Hughes, J. W. LeDuc, W. J. Bellini, and L. J. Anderson. 2003.** A novel coronavirus associated with severe acute respiratory syndrome. *N.Engl.J.Med.* **348**:1953-1966
96. **Kummerer, B. M. and C. M. Rice. 2002.** Mutations in the yellow fever virus nonstructural protein NS2A selectively block production of infectious particles. *J.Virol.* **76**:4773-4784
97. **Kushner, D. B., B. D. Lindenbach, V. Z. Grdzlishvili, A. O. Noueiry, S. M. Paul, and P. Ahlquist. 2003.** Systematic, genome-wide identification of host genes affecting replication of a positive-strand RNA virus. *Proc.Natl.Acad.Sci.U.S.A* **100**:15764-15769
98. **Kuyumcu-Martinez, N. M., M. Joachims, and R. E. Lloyd. 2002.** Efficient cleavage of ribosome-associated poly(A)-binding protein by enterovirus 3C protease. *J.Virol.* **76**:2062-2074
99. **Lackner, T., A. Muller, A. Pankraz, P. Becher, H. J. Thiel, A. E. Gorbalenya, and N. Tautz. 2004.** Temporal modulation of an autoprotease is crucial for replication and pathogenicity of an RNA virus. *J.Virol.* **78**:10765-10775
100. **Larsen, G. R., A. J. Dorner, T. J. Harris, and E. Wimmer. 1980.** The structure of poliovirus replicative form. *Nucleic Acids Res.* **8**:1217-1229
101. **Lee, Y. F., A. Nomoto, B. M. Detjen, and E. Wimmer. 1977.** A protein covalently linked to poliovirus genome RNA. *Proc.Natl.Acad.Sci.U.S.A* **74**:59-63
102. **Leung, J. Y., G. P. Pijlman, N. Kondratieva, J. Hyde, J. M. Mackenzie, and A. A. Khromykh. 2008.** Role of nonstructural protein NS2A in flavivirus assembly. *J.Virol.* **82**:4731-4741
103. **Li, L., H. Kang, P. Liu, N. Makkinje, S. T. Williamson, J. L. Leibowitz, and D. P. Giedroc. 2008.** Structural lability in stem-loop 1 drives a 5'UTR-3'UTR interaction in coronavirus replication. *J.Mol. Biol.* **377**:790-803
104. **Lin, H. X. and K. A. White. 2004.** A complex network of RNA-RNA interactions controls subgenomic mRNA transcription in a tombusvirus. *EMBO J.* **23**:3365-3374
105. **Lin, H. X., W. Xu, and K. A. White. 2007.** A multicomponent RNA-based control system regulates subgenomic mRNA transcription in a tombusvirus. *J.Virol.* **81**:2429-2439
106. **Lindenbach, B. D., M. J. Evans, A. J. Syder, B. Wolk, T. L. Tellinghuisen, C. C. Liu, T. Maruyama, R. O. Hynes, D. R. Burton, J. A. McKeating, and C. M. Rice. 2005.** Complete replication of hepatitis C virus in cell culture. *Science* **309**:623-626



107. **Lindenbach, B. D. and C. M. Rice. 2003.** Molecular biology of flaviviruses. *Adv.Virus Res.* **59**:23-61
108. **Liu, W. J., H. B. Chen, and A. A. Khromykh. 2003.** Molecular and functional analyses of Kunjin virus infectious cDNA clones demonstrate the essential roles for NS2A in virus assembly and for a nonconservative residue in NS3 in RNA replication. *J.Virol.* **77**:7804-7813
109. **Liu, W. J., P. L. Sedlak, N. Kondratieva, and A. A. Khromykh. 2002.** Complementation analysis of the flavivirus Kunjin NS3 and NS5 proteins defines the minimal regions essential for formation of a replication complex and shows a requirement of NS3 in cis for virus assembly. *J.Virol.* **76**:10766-10775
110. **Lobigs, M. 1993.** Flavivirus premembrane protein cleavage and spike heterodimer secretion require the function of the viral proteinase NS3. *Proc.Natl.Acad.Sci.U.S.A* **90**:6218-6222
111. **Lohmann, V., F. Korner, U. Herian, and R. Bartenschlager. 1997.** Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. *J.Virol.* **71**:8416-8428
112. **Lohmann, V., F. Korner, J. Koch, U. Herian, L. Theilmann, and R. Bartenschlager. 1999.** Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* **285**:110-113
113. **Lorenz, I. C., J. Marcotrigiano, T. G. Dentzer, and C. M. Rice. 2006.** Structure of the catalytic domain of the hepatitis C virus NS2-3 protease. *Nature* **442**:831-835
114. **Ma, Y., J. Yates, Y. Liang, S. M. Lemon, and M. Yi. 2008.** NS3 helicase domains involved in infectious intracellular hepatitis C virus particle assembly. *J.Virol.* **82**:7624-7639
115. **Mackenzie, J. 2005.** Wrapping things up about virus RNA replication. *Traffic.* **6**:967-977
116. **Mackenzie, J. M., M. K. Jones, and P. R. Young. 1996.** Improved membrane preservation of flavivirus-infected cells with cryosectioning. *J.Virol.Methods* **56**:67-75
117. **Mackenzie, J. M., A. A. Khromykh, M. K. Jones, and E. G. Westaway. 1998.** Subcellular localization and some biochemical properties of the flavivirus Kunjin nonstructural proteins NS2A and NS4A. *Virology* **245**:203-215
118. **Mackenzie, J. M. and E. G. Westaway. 2001.** Assembly and maturation of the flavivirus Kunjin virus appear to occur in the rough endoplasmic reticulum and along the secretory pathway, respectively. *J.Virol.* **75**:10787-10799
119. **Makino, S., K. Yokomori, and M. M. Lai. 1990.** Analysis of efficiently packaged defective interfering RNAs of murine coronavirus: localization of a possible RNA-packaging signal. *J.Virol.* **64**:6045-6053
120. **Mas, A., I. Alves-Rodrigues, A. Noueiry, P. Ahlquist, and J. Diez. 2006.** Host deadenylation-dependent mRNA decapping factors are required for a key step in brome mosaic virus RNA replication. *J.Virol.* **80**:246-251
121. **Masaki, T., R. Suzuki, K. Murakami, H. Aizaki, K. Ishii, A. Murayama, T. Date, Y. Matsuura, T. Miyamura, T. Wakita, and T. Suzuki. 2008.** Interaction of hepatitis C virus nonstructural protein 5A with core protein is critical for the production of infectious virus particles. *J.Virol.* **82**:7964-7976
122. **McCartney, A. W., J. S. Greenwood, M. R. Fabian, K. A. White, and R. T. Mullen. 2005.** Localization of the tomato bushy stunt virus replication protein p33 reveals a peroxisome-to-endoplasmic reticulum sorting pathway. *Plant Cell* **17**:3513-3531
123. **Miller, R. H. and R. H. Purcell. 1990.** Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups. *Proc.Natl.Acad.Sci.U.S.A* **87**:2057-2061
124. **Miller, S., S. Kastner, J. Krijnse-Locker, S. Buhler, and R. Bartenschlager. 2007.** The non-structural protein 4A of dengue virus is an integral membrane protein inducing membrane alterations in a 2K-regulated manner. *J.Biol.Chem.* **282**:8873-8882



125. **Miller, S. and J. Krijnse-Locker. 2008.** Modification of intracellular membrane structures for virus replication. *Nat Rev.Microbiol.* **6**:363-374
126. **Miller, W. A. and G. Koev. 2000.** Synthesis of subgenomic RNAs by positive-strand RNA viruses. *Virology* **273**:1-8
127. **Minskaia, E., T. Hertzog, A. E. Gorbalenya, V. Campanacci, C. Cambillau, B. Canard, and J. Ziebuhr. 2006.** Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc.Natl.Acad.Sci.U.S.A* **103**:5108-5113
128. **Miyanari, Y., K. Atsuzawa, N. Usuda, K. Watashi, T. Hishiki, M. Zayas, R. Bartenschlager, T. Wakita, M. Hijikata, and K. Shimotohno. 2007.** The lipid droplet is an important organelle for hepatitis C virus production. *Nat.Cell Biol.* **9**:1089-1097
129. **Molenkamp, R., B. C. D. Rozier, S. Greve, W. J. M. Spaan, and E. J. Snijder. 2000.** Isolation and characterization of an arterivirus defective interfering RNA genome. *J. Virol.* **74**:3156-3165
130. **Molenkamp, R., H. van Tol, B. C. D. Rozier, Y. van der Meer, W. J. M. Spaan, and E. J. Snijder. 2000.** The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription. *J.Gen.Virol* **81**:2491-2496
131. **Molla, A., A. V. Paul, and E. Wimmer. 1991.** Cell-free, de novo synthesis of poliovirus. *Science* **254**:1647-1651
132. **Monkewich, S., H. X. Lin, M. R. Fabian, W. Xu, H. Na, D. Ray, O. A. Chernysheva, P. D. Nagy, and K. A. White. 2005.** The p92 polymerase coding region contains an internal RNA element required at an early step in Tombusvirus genome replication. *J.Virol.* **79**:4848-4858
133. **Moradpour, D., F. Penin, and C. M. Rice. 2007.** Replication of hepatitis C virus. *Nat.Rev.Microbiol.* **5**:453-463
134. **Morasco, B. J., N. Sharma, J. Parilla, and J. B. Flanagan. 2003.** Poliovirus cre(2C)-dependent synthesis of VPgpUpU is required for positive- but not negative-strand RNA synthesis. *J.Virol.* **77**:5136-5144
135. **Moreno, J. L., S. Zuniga, L. Enjuanes, and I. Sola. 2008.** Identification of a coronavirus transcription enhancer. *J.Virol.* **82**:3882-3893
136. **Mosimann, S. C., M. M. Cherney, S. Sia, S. Plotch, and M. N. James. 1997.** Refined X-ray crystallographic structure of the poliovirus 3C gene product. *J.Mol.Biol.* **273**:1032-1047
137. **Murray, C. L., C. T. Jones, and C. M. Rice. 2008.** Architects of assembly: roles of Flaviviridae non-structural proteins in virion morphogenesis. *Nat.Rev.Microbiol.* **6**:699-708
138. **Murray, K. E. and D. J. Barton. 2003.** Poliovirus CRE-dependent VPg uridylylation is required for positive-strand RNA synthesis but not for negative-strand RNA synthesis. *J.Virol.* **77**:4739-4750
139. **Murray, K. E., B. P. Steil, A. W. Roberts, and D. J. Barton. 2004.** Replication of poliovirus RNA with complete internal ribosome entry site deletions. *J.Virol.* **78**:1393-1402
140. **Narayanan, K. and S. Makino. 2001.** Cooperation of an RNA packaging signal and a viral envelope protein in coronavirus RNA packaging. *J.Virol.* **75**:9059-9067
141. **Nestorowicz, A., T. J. Chambers, and C. M. Rice. 1994.** Mutagenesis of the yellow fever virus NS2A/2B cleavage site: effects on proteolytic processing, viral replication, and evidence for alternative processing of the NS2A protein. *Virology* **199**:114-123
142. **Neuman, B. W., B. D. Adair, M. Yeager, and M. J. Buchmeier. 2008.** Purification and electron cryomicroscopy of coronavirus particles. *Methods Mol.Biol.* **454**:129-136
143. **Nomoto, A., Y. F. Lee, and E. Wimmer. 1976.** The 5' end of poliovirus mRNA is not capped with m7G(5')ppp(5')Np. *Proc.Natl.Acad.Sci.U.S.A* **73**:375-380
144. **Novak, J. E. and K. Kirkegaard. 1994.** Coupling between genome translation and replication in an RNA virus. *Genes Dev.* **8**:1726-1737



145. **Novoa, R. R., G. Calderita, R. Arranz, J. Fontana, H. Granzow, and C. Risco. 2005.** Virus factories: associations of cell organelles for viral replication and morphogenesis. *Biol.Cell* **97**:147-172
146. **Nugent, C. I., K. L. Johnson, P. Sarnow, and K. Kirkegaard. 1999.** Functional coupling between replication and packaging of poliovirus replicon RNA. *J.Virol.* **73**:427-435
147. **Oh, H. S., H. B. Pathak, I. G. Goodfellow, J. J. Arnold, and C. E. Cameron. 2009.** Insight into poliovirus genome replication and encapsidation obtained from studies of 3B-3C cleavage site mutants. *J.Virol.* **83**:9370-9387
148. **Oostra, M., M. C. Hagemeijer, G. M. van, C. P. Bekker, E. G. te Lintelo, P. J. Rottier, and C. A. de Haan. 2008.** Topology and membrane anchoring of the coronavirus replication complex: not all hydrophobic domains of nsp3 and nsp6 are membrane spanning. *J.Virol.* **82**:12392-12405
149. **Oostra, M., E. G. te Lintelo, M. Deijs, M. H. Verheije, P. J. Rottier, and C. A. de Haan. 2007.** Localization and membrane topology of coronavirus nonstructural protein 4: involvement of the early secretory pathway in replication. *J.Virol.* **81**:12323-12336
150. **Pan, J., X. Peng, Y. Gao, Z. Li, X. Lu, Y. Chen, M. Ishaq, D. Liu, M. L. Dediego, L. Enjuanes, and D. Guo. 2008.** Genome-wide analysis of protein-protein interactions and involvement of viral proteins in SARS-CoV replication. *PLoS.One.* **3**:e3299
151. **Panavas, T., E. Serviene, J. Brasher, and P. D. Nagy. 2005.** Yeast genome-wide screen reveals dissimilar sets of host genes affecting replication of RNA viruses. *Proc.Natl.Acad.Sci.U.S.A* **102**:7326-7331
152. **Panaviene, Z., J. M. Baker, and P. D. Nagy. 2003.** The overlapping RNA-binding domains of p33 and p92 replicase proteins are essential for tombusvirus replication. *Virology* **308**:191-205
153. **Panaviene, Z., T. Panavas, S. Serva, and P. D. Nagy. 2004.** Purification of the cucumber necrosis virus replicase from yeast cells: role of coexpressed viral RNA in stimulation of replicase activity. *J.Virol.* **78**:8254-8263
154. **Pasternak, A. O., W. J. Spaan, and E. J. Snijder. 2006.** Nidovirus transcription: how to make sense...? *J.Gen.Virol.* **87**:1403-1421
155. **Pasternak, A. O., W. J. M. Spaan, and E. J. Snijder. 2004.** Regulation of relative abundance of Arterivirus subgenomic rnrRNAs. *Journal of Virology* **78**:8102-8113
156. **Pasternak, A. O., E. van den Born, W. J. M. Spaan, and E. J. Snijder. 2001.** Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. *EMBO J.* **20**:7220-7228
157. **Pasternak, A. O., E. van den Born, W. J. M. Spaan, and E. J. Snijder. 2003.** The stability of the duplex between sense and antisense transcription-regulating sequences is a crucial factor in arterivirus subgenomic mRNA synthesis. *J. Virol.* **77**:1175-1183
158. **Patkar, C. G. and R. J. Kuhn. 2008.** Yellow Fever virus NS3 plays an essential role in virus assembly independent of its known enzymatic functions. *J.Virol.* **82**:3342-3352
159. **Pedersen, K. W., Y. van der Meer, N. Roos, and E. J. Snijder. 1999.** 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.***73**:2016-2026
160. **Perera, R., S. Daijogo, B. L. Walter, J. H. Nguyen, and B. L. Semler. 2007.** Cellular protein modification by poliovirus: the two faces of poly(rC)-binding protein. *J.Virol.* **81**:8919-8932
161. **Pietschmann, T., A. Kaul, G. Koutsoudakis, A. Shavinskaya, S. Kallis, E. Steinmann, K. Abid, F. Negro, M. Dreux, F. L. Cosset, and R. Bartenschlager. 2006.** Construction and characterization of infectious intragenotypic and intergenotypic hepatitis C virus chimeras. *Proc.Natl.Acad. Sci.U.S.A* **103**:7408-7413
162. **Pietschmann, T., M. Zayas, P. Meuleman, G. Long, N. Appel, G. Koutsoudakis, S. Kallis, G. Leroux-Roels, V. Lohmann, and R. Bartenschlager. 2009.** Production of infectious genotype 1b

- virus particles in cell culture and impairment by replication enhancing mutations. *PLoS.Pathog.* **5**:e1000475
163. **Pijlman, G. P., N. Kondratieva, and A. A. Khromykh. 2006.** Translation of the flavivirus kunjin NS3 gene in cis but not its RNA sequence or secondary structure is essential for efficient RNA packaging. *J.Virol.* **80**:11255-11264
  164. **Pogany, J. and P. D. Nagy. 2008.** Authentic replication and recombination of Tomato bushy stunt virus RNA in a cell-free extract from yeast. *J.Virol.* **82**:5967-5980
  165. **Pogany, J., J. Stork, Z. Li, and P. D. Nagy. 2008.** In vitro assembly of the Tomato bushy stunt virus replicase requires the host Heat shock protein 70. *Proc.Natl.Acad.Sci.U.S.A* **105**:19956-19961
  166. **Pogany, J., K. A. White, and P. D. Nagy. 2005.** Specific binding of tombusvirus replication protein p33 to an internal replication element in the viral RNA is essential for replication. *J.Virol.* **79**:4859-4869
  167. **Posthuma, C. C., K. W. Pedersen, Z. Lu, R. G. Joosten, N. Roos, J. C. Zevenhoven-Dobbe, and E. J. Snijder. 2008.** Formation of the arterivirus replication/transcription complex: a key role for nonstructural protein 3 in the remodeling of intracellular membranes. *J.Virol.* **82**:4480-4491
  168. **Quinkert, D., R. Bartenschlager, and V. Lohmann. 2005.** Quantitative analysis of the hepatitis C virus replication complex. *J.Virol.* **79**:13594-13605
  169. **Rajendran, K. S. and P. D. Nagy. 2003.** Characterization of the RNA-binding domains in the replicase proteins of tomato bushy stunt virus. *J.Virol.* **77**:9244-9258
  170. **Romero-Lopez, C. and A. Berzal-Herranz. 2009.** A long-range RNA-RNA interaction between the 5' and 3' ends of the HCV genome. *RNA.* **15**:1740-1752
  171. **Russo, M., J. Burgyan, and G. P. Martelli. 1994.** Molecular biology of tombusviridae. *Adv.Virus Res.* **44**:381-428
  172. **Sachs, A. B. and S. Buratowski. 1997.** Common themes in translational and transcriptional regulation. *Trends Biochem.Sci.* **22**:189-192
  173. **Samsa, M. M., J. A. Mondotte, N. G. Iglesias, I. Assuncao-Miranda, G. Barbosa-Lima, A. T. Da Poian, P. T. Bozza, and A. V. Gamarnik. 2009.** Dengue virus capsid protein usurps lipid droplets for viral particle formation. *PLoS.Pathog.* **5**:e1000632
  174. **Sawicki, D. L., T. Wang, and S. G. Sawicki. 2001.** The RNA structures engaged in replication and transcription of the A59 strain of mouse hepatitis virus. *J.Gen.Virol.* **82**:385-396
  175. **Sawicki, S. G. and D. L. Sawicki. 1995.** Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. *Adv.Exp.Med.Biol.* **380**:499-506
  176. **Sawicki, S. G., D. L. Sawicki, and S. G. Siddell. 2007.** A contemporary view of coronavirus transcription. *J.Virol.* **81**:20-29
  177. **Sawicki, S. G., D. L. Sawicki, D. Younker, Y. Meyer, V. Thiel, H. Stokes, and S. G. Siddell. 2005.** Functional and genetic analysis of coronavirus replicase-transcriptase proteins. *PLoS.Pathog.* **1**:e39
  178. **Schalich, J., S. L. Allison, K. Stiasny, C. W. Mandl, C. Kunz, and F. X. Heinz. 1996.** Recombinant subviral particles from tick-borne encephalitis virus are fusogenic and provide a model system for studying flavivirus envelope glycoprotein functions. *J.Virol.* **70**:4549-4557
  179. **Schaller, T., N. Appel, G. Koutsoudakis, S. Kallis, V. Lohmann, T. Pietschmann, and R. Bartenschlager. 2007.** Analysis of hepatitis C virus superinfection exclusion by using novel fluorochrome gene-tagged viral genomes. *J.Virol.* **81**:4591-4603
  180. **Schelle, B., N. Karl, B. Ludewig, S. G. Siddell, and V. Thiel. 2005.** Selective replication of coronavirus genomes that express nucleocapsid protein. *J.Virol.* **79**:6620-6630



181. **Scholthof, H. B., B. Desvoyes, J. Kuecker, and E. Whitehead. 1999.** Biological activity of two tombusvirus proteins translated from nested genes is influenced by dosage control via context-dependent leaky scanning. *Mol. Plant-Microbe Interact.* **12**:670-679
182. **Scholthof, K. B., H. B. Scholthof, and A. O. Jackson. 1995.** The tomato bushy stunt virus replicase proteins are coordinately expressed and membrane associated. *Virology* **208**:365-369
183. **Schutze, H., R. Ulferts, B. Schelle, S. Bayer, H. Granzow, B. Hoffmann, T. C. Mettenleiter, and J. Ziebuhr. 2006.** Characterization of White breem virus reveals a novel genetic cluster of nidoviruses. *J.Virol.* **80**:11598-11609
184. **Serrano, P., M. R. Pulido, M. Saiz, and E. Martinez-Salas. 2006.** The 3' end of the foot-and-mouth disease virus genome establishes two distinct long-range RNA-RNA interactions with the 5' end region. *J.Gen.Virol.* **87**:3013-3022
185. **Serviene, E., Y. Jiang, C. P. Cheng, J. Baker, and P. D. Nagy. 2006.** Screening of the yeast yTHC collection identifies essential host factors affecting tombusvirus RNA recombination. *J.Virol.* **80**:1231-1241
186. **Serviene, E., N. Shapka, C. P. Cheng, T. Panavas, B. Phuangrat, J. Baker, and P. D. Nagy. 2005.** Genome-wide screen identifies host genes affecting viral RNA recombination. *Proc.Natl.Acad. Sci.U.S.A* **102**:10545-10550
187. **Seybert, A., C. C. Posthuma, L. C. van Dinten, E. J. Snijder, A. E. Gorbalenya, and J. Ziebuhr. 2005.** A complex zinc finger controls the enzymatic activities of nidovirus helicases. *J.Virol.* **79**:696-704
188. **Seybert, A., L. C. van Dinten, E. J. Snijder, and J. Ziebuhr. 2000.** Biochemical characterization of the equine arteritis virus helicase suggests a close functional relationship between arterivirus and coronavirus helicases. *J. Virol.* **74**:9586-9593
189. **Shapka, N., J. Stork, and P. D. Nagy. 2005.** Phosphorylation of the p33 replication protein of Cucumber necrosis tombusvirus adjacent to the RNA binding site affects viral RNA replication. *Virology* **343**:65-78
190. **Shavinskaya, A., S. Boulant, F. Penin, J. McLauchlan, and R. Bartenschlager. 2007.** The lipid droplet binding domain of hepatitis C virus core protein is a major determinant for efficient virus assembly. *J.Biol.Chem.* **282**:37158-37169
191. **Shen, X. and P. S. Masters. 2001.** Evaluation of the role of heterogeneous nuclear ribonucleoprotein A1 as a host factor in murine coronavirus discontinuous transcription and genome replication. *Proc.Natl.Acad.Sci.U.S.A* **98**:2717-2722
192. **Shi, S. T., P. Huang, H. P. Li, and M. M. Lai. 2000.** Heterogeneous nuclear ribonucleoprotein A1 regulates RNA synthesis of a cytoplasmic virus. *EMBO J.* **19**:4701-4711
193. **Shi, S. T., G. Y. Yu, and M. M. Lai. 2003.** Multiple type A/B heterogeneous nuclear ribonucleoproteins (hnRNPs) can replace hnRNP A1 in mouse hepatitis virus RNA synthesis. *J.Virol.* **77**:10584-10593
194. **Silvera, D., A. V. Gamarnik, and R. Andino. 1999.** The N-terminal K homology domain of the poly(rC)-binding protein is a major determinant for binding to the poliovirus 5'-untranslated region and acts as an inhibitor of viral translation. *J.Biol.Chem.* **274**:38163-38170
195. **Sittidilokratna, N., S. Dangtip, J. A. Cowley, and P. J. Walker. 2008.** RNA transcription analysis and completion of the genome sequence of yellow head nidovirus. *Virus Res.* **136**:157-165
196. **Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. M. Poon, Y. Guan, M. Rozanov, W. J. M. Spaan, and A. E. Gorbalenya. 2003.** Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J. Mol. Biol.* **331**:991-1004

197. **Snijder, E. J., J. A. den Boon, M. C. Horzinek, and W. J. Spaan. 1991.** Comparison of the genome organization of toro- and coronaviruses: evidence for two nonhomologous RNA recombination events during Berne virus evolution. *Virology* **180**:448-452
198. **Snijder, E. J., H. van Tol, N. Roos, and K. W. Pedersen. 2001.** Non-structural proteins 2 and 3 interact to modify host cell membranes during the formation of the arterivirus replication complex. *J Gen Virol* **82**:985-994
199. **Snijder, E. J., A. L. Wassenaar, and W. J. Spaan. 1994.** Proteolytic processing of the replicase ORF1a protein of equine arteritis virus. *J.Virol.* **68**:5755-5764
200. **Sola, I., J. L. Moreno, S. Zuniga, S. Alonso, and L. Enjuanes. 2005.** Role of nucleotides immediately flanking the transcription-regulating sequence core in coronavirus subgenomic mRNA synthesis. *J.Virol.* **79**:2506-2516
201. **Spagnolo, J. F. and B. G. Hogue. 2000.** Host protein interactions with the 3' end of bovine coronavirus RNA and the requirement of the poly(A) tail for coronavirus defective genome replication. *J Virol* **74**:5053-5065
202. **Steinmann, E., F. Penin, S. Kallis, A. H. Patel, R. Bartenschlager, and T. Pietschmann. 2007.** Hepatitis C virus p7 protein is crucial for assembly and release of infectious virions. *PLoS.Pathog.* **3**:e103
203. **Svitkin, Y. V., M. Costa-Mattioli, B. Herdy, S. Perreault, and N. Sonenberg. 2007.** Stimulation of picornavirus replication by the poly(A) tail in a cell-free extract is largely independent of the poly(A) binding protein (PABP). *RNA* **13**:2330-2340
204. **Tan, B. H., J. Fu, R. J. Sugrue, E. H. Yap, Y. C. Chan, and Y. H. Tan. 1996.** Recombinant dengue type 1 virus NS5 protein expressed in *Escherichia coli* exhibits RNA-dependent RNA polymerase activity. *Virology* **216**:317-325
205. **Tanji, Y., T. Kaneko, S. Satoh, and K. Shimotohno. 1995.** Phosphorylation of hepatitis C virus-encoded nonstructural protein NS5A. *J.Virol.* **69**:3980-3986
206. **te Velthuis, A. J., J. J. Arnold, C. E. Cameron, S. H. van den Worm, and E. J. Snijder. 2010.** The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. *Nucleic Acids Res.* **38**:203-214
207. **Tellinghuisen, T. L., K. L. Foss, and J. Treadaway. 2008.** Regulation of hepatitis C virion production via phosphorylation of the NS5A protein. *PLoS.Pathog.* **4**:e1000032
208. **Tellinghuisen, T. L., J. Marcotrigiano, and C. M. Rice. 2005.** Structure of the zinc-binding domain of an essential component of the hepatitis C virus replicase. *Nature* **435**:374-379
209. **Thiel, V., J. Herold, B. Schelle, and S. G. Siddell. 2001.** Viral replicase gene products suffice for coronavirus discontinuous transcription. *J.Virol.* **75**:6676-6681
210. **Thurner, C., C. Witwer, I. L. Hofacker, and P. F. Stadler. 2004.** Conserved RNA secondary structures in Flaviviridae genomes. *J.Gen.Virol.* **85**:1113-1124
211. **Tijms, M. A., Y. van der Meer, and E. J. Snijder. 2002.** Nuclear localization of non-structural protein 1 and nucleocapsid protein of equine arteritis virus. *J.Gen.Virol.* **83**:795-800
212. **Tijms, M. A., L. C. van Dinten, A. E. Gorbalenya, and E. J. Snijder. 2001.** A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus. *Proc.Natl.Acad.Sci.U.S.A* **98**:1889-1894
213. **van Aken, D., J. Zevenhoven-Dobbe, A. E. Gorbalenya, and E. J. Snijder. 2006.** Proteolytic maturation of replicase polyprotein pp1a by the nsp4 main proteinase is essential for equine arteritis virus replication and includes internal cleavage of nsp7. *J.Gen.Virol.* **87**:3473-3482
214. **van den Born, E., A. P. Gultyaev, and E. J. Snijder. 2004.** Secondary structure and function of the 5'-proximal region of the equine arteritis virus RNA genome. *RNA* **10**:424-437



215. **van den Born, E., C. C. Posthuma, A. P. Gultyaev, and E. J. Snijder. 2005.** Discontinuous subgenomic RNA synthesis in arteriviruses is guided by an RNA hairpin structure located in the genomic leader region. *J. Virol.* **79**:6312-6324
216. **van der Meer, Y., E. J. Snijder, J. C. Dobbe, S. Schleich, M. R. Denison, W. J. Spaan, and J. K. Locker. 1999.** Localization of mouse hepatitis virus nonstructural proteins and RNA synthesis indicates a role for late endosomes in viral replication. *J. Virol.* **73**:7641-7657
217. **van Dinten, L. C., J. A. den Boon, A. L. M. Wassenaar, W. J. M. Spaan, and E. J. Snijder. 1997.** An infectious arterivirus cDNA clone: Identification of a replicase point mutation that abolishes discontinuous mRNA transcription. *Proc. Natl. Acad. Sci. U.S.A.* **94**:991-996
218. **van Dinten, L. C., S. Rensen, A. E. Gorbalenya, and E. J. Snijder. 1999.** Proteolytic processing of the open reading frame 1b-encoded part of arterivirus replicase is mediated by nsp4 serine protease and is essential for virus replication. *J. Virol.* **73**:2027-2037
219. **van Dinten, L. C., H. van Tol, A. E. Gorbalenya, and E. J. Snijder. 2000.** The predicted metal-binding region of the arterivirus helicase protein is involved in subgenomic mRNA synthesis, genome replication, and virion biogenesis. *J. Virol.* **74**:5213-5223
220. **van Dinten, L. C., A. L. Wassenaar, A. E. Gorbalenya, W. J. Spaan, and E. J. Snijder. 1996.** Processing of the equine arteritis virus replicase ORF1b protein: identification of cleavage products containing the putative viral polymerase and helicase domains. *J. Virol.* **70**:6625-6633
221. **van Dyke, T. A. and J. B. Flanagan. 1980.** Identification of poliovirus polypeptide P63 as a soluble RNA-dependent RNA polymerase. *J. Virol.* **35**:732-740
222. **van Hemert, M. J., A. H. de Wilde, A. E. Gorbalenya, and E. J. Snijder. 2008.** The in vitro RNA synthesizing activity of the isolated arterivirus replication/transcription complex is dependent on a host factor. *J. Biol. Chem.* **283**:16525-16536
223. **van Hemert, M. J., S. H. van den Worm, K. Knoops, A. M. Mommaas, A. E. Gorbalenya, and E. J. Snijder. 2008.** SARS-coronavirus replication/transcription complexes are membrane-protected and need a host factor for activity in vitro. *PLoS Pathog.* **4**:e1000054
224. **van Marle, G., J. C. Dobbe, A. P. Gultyaev, W. Luytjes, W. J. M. Spaan, and E. J. Snijder. 1999.** Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and anti-sense transcription-regulating sequences. *Proc. Natl. Acad. Sci. U.S.A.* **96**:12056-12061
225. **van Marle, G., L. C. van Dinten, W. J. M. Spaan, W. Luytjes, and E. J. Snijder. 1999.** Characterization of an equine arteritis virus replicase mutant defective in subgenomic mRNA synthesis. *J. Virol.* **73**:5274-5281
226. **van Vliet, A. L., S. L. Smits, P. J. Rottier, and R. J. de Groot. 2002.** Discontinuous and non-discontinuous subgenomic RNA transcription in a nidovirus. *EMBO J.* **21**:6571-6580
227. **Varnavski, A. N., P. R. Young, and A. A. Khromykh. 2000.** Stable high-level expression of heterologous genes in vitro and in vivo by noncytopathic DNA-based Kunjin virus replicon vectors. *J. Virol.* **74**:4394-4403
228. **Vennema, H., G. J. Godeke, J. W. Rossen, W. F. Voorhout, M. C. Horzinek, D. J. Opstelten, and P. J. Rottier. 1996.** Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. *EMBO J.* **15**:2020-2028
229. **Verheije, M. H., R. C. Olsthoorn, M. V. Kroese, P. J. Rottier, and J. J. Meulenberg. 2002.** Kissing interaction between 3' noncoding and coding sequences is essential for porcine arterivirus RNA replication. *J. Virol.* **76**:1521-1526
230. **von Brunn A., C. Teepe, J. C. Simpson, R. Pepperkok, C. C. Friedel, R. Zimmer, R. Roberts, R. Baric, and J. Haas. 2007.** Analysis of intraviral protein-protein interactions of the SARS coronavirus ORFome. *PLoS One.* **2**:e459

231. **Wakita, T., T. Pietschmann, T. Kato, T. Date, M. Miyamoto, Z. Zhao, K. Murthy, A. Habermann, H. G. Krausslich, M. Mizokami, R. Bartenschlager, and T. J. Liang. 2005.** Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat.Med.* **11**:791-796
232. **Wang, R. Y. and P. D. Nagy. 2008.** Tomato bushy stunt virus co-opts the RNA-binding function of a host metabolic enzyme for viral genomic RNA synthesis. *Cell Host&Microbe* **3**:178-187
233. **Wang, S., L. Mortazavi, and K. A. White. 2008.** Higher-order RNA structural requirements and small-molecule induction of tombusvirus subgenomic mRNA transcription. *J.Virol.* **82**:3864-3871
234. **Wang, Z., N. Day, P. Trifillis, and M. Kiledjian. 1999.** An mRNA stability complex functions with poly(A)-binding protein to stabilize mRNA in vitro. *Mol.Cell Biol.* **19**:4552-4560
235. **Wassenaar, A. L. M., W. J. M. Spaan, A. E. Gorbalenya, and E. J. Snijder. 1997.** Alternative proteolytic processing of the arterivirus replicase ORF1a polyprotein: evidence that NSP2 acts as a cofactor for the NSP4 serine protease. *J. Virol.* **71**:9313-9322
236. **Welbourn, S., R. Green, I. Gamache, S. Dandache, V. Lohmann, R. Bartenschlager, K. Meero-vitch, and A. Pause. 2005.** Hepatitis C virus NS2/3 processing is required for NS3 stability and viral RNA replication. *J.Biol.Chem.* **280**:29604-29611
237. **Welbourn, S., V. Jirasko, V. Breton, S. Reiss, F. Penin, R. Bartenschlager, and A. Pause. 2009.** Investigation of a role for lysine residues in non-structural proteins 2 and 2/3 of the hepatitis C virus for their degradation and virus assembly. *J.Gen.Virol.* **90**:1071-1080
238. **Welsch, S., S. Miller, I. Romero-Brey, A. Merz, C. K. Bleck, P. Walther, S. D. Fuller, C. Antony, J. Krijnse-Locker, and R. Bartenschlager. 2009.** Composition and three-dimensional architecture of the dengue virus replication and assembly sites. *Cell Host&Microbe* **5**:365-375
239. **Wengler, G. and G. Wengler. 1991.** The carboxy-terminal part of the NS 3 protein of the West Nile flavivirus can be isolated as a soluble protein after proteolytic cleavage and represents an RNA-stimulated NTPase. *Virology* **184**:707-715
240. **White, K. A. 2002.** The premature termination model: a possible third mechanism for subgenomic mRNA transcription in (+)-strand RNA viruses. *Virology* **304**:147-154
241. **Williams, G. D., R. Y. Chang, and D. A. Brian. 1999.** A phylogenetically conserved hairpin-type 3' untranslated region pseudoknot functions in coronavirus RNA replication. *J.Virol.* **73**:8349-8355
242. **Woo, K., M. Joo, K. Narayanan, K. H. Kim, and S. Makino. 1997.** Murine coronavirus packaging signal confers packaging to nonviral RNA. *J.Virol.* **71**:824-827
243. **Wu, B., J. Pogany, H. Na, B. L. Nicholson, P. D. Nagy, and K. A. White. 2009.** A discontinuous RNA platform mediates RNA virus replication: building an integrated model for RNA-based regulation of viral processes. *PLoS.Pathog.* **5**:e1000323
244. **Wu, B. and K. A. White. 2007.** Uncoupling RNA virus replication from transcription via the polymerase: functional and evolutionary insights. *EMBO J.* **26**:5120-5130
245. **Xu, W. and K. A. White. 2008.** Subgenomic mRNA transcription in an aureusvirus: down-regulation of transcription and evolution of regulatory RNA elements. *Virology* **371**:430-438
246. **Yakobson, E., A. Mikhejeva, and Y. Z. Ghendon. 1973.** On the kinetics of poliovirus-specific protein and RNA synthesis in poliovirus-infected cells. *Arch.Gesamte Virusforsch.* **40**:161-163
247. **Yi, M., Y. Ma, J. Yates, and S. M. Lemon. 2007.** Compensatory mutations in E1, p7, NS2, and NS3 enhance yields of cell culture-infectious intergenotypic chimeric hepatitis C virus. *J.Virol.* **81**:629-638
248. **Yi, M., Y. Ma, J. Yates, and S. M. Lemon. 2009.** Trans-complementation of an NS2 defect in a late step in hepatitis C virus (HCV) particle assembly and maturation. *PLoS.Pathog.* **5**:e1000403
249. **Yogo, Y. and E. Wimmer. 1972.** Polyadenylic acid at the 3'-terminus of poliovirus RNA. *Proc.Natl. Acad.Sci.U.S.A* **69**:1877-1882



250. **Yount, B., K. M. Curtis, E. A. Fritz, L. E. Hensley, P. B. Jahrling, E. Prentice, M. R. Denison, T. W. Geisbert, and R. S. Baric. 2003.** Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. *Proc.Natl.Acad.Sci.U.S.A* **100**:12995-13000
251. **Yount, B., M. R. Denison, S. R. Weiss, and R. S. Baric. 2002.** Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. *J.Virol.* **76**:11065-11078
252. **Zhang, G., V. Slowinski, and K. A. White. 1999.** Subgenomic mRNA regulation by a distal RNA element in a (+)-strand RNA virus. *RNA* **5**:550-561
253. **Zhong, J., P. Gastaminza, G. Cheng, S. Kapadia, T. Kato, D. R. Burton, S. F. Wieland, S. L. Up-  
richard, T. Wakita, and F. V. Chisari. 2005.** Robust hepatitis C virus infection in vitro. *Proc.Natl.  
Acad.Sci.U.S.A* **102**:9294-9299
254. **Ziebuhr, J. and S. G. Siddell. 1999.** Processing of the human coronavirus 229E replicase polypro-  
teins by the virus-encoded 3C-like proteinase: identification of proteolytic products and cleavage  
sites common to pp1a and pp1ab. *J.Virol.* **73**:177-185
255. **Ziebuhr, J., E. J. Snijder, and A. E. Gorbalenya. 2000.** Virus-encoded proteinases and proteolytic  
processing in the Nidovirales. *J.Gen.Virol.* **81**:853-879
256. **Zuniga, S., I. Sola, S. Alonso, and L. Enjuanes. 2004.** Sequence motifs involved in the regulation  
of discontinuous coronavirus subgenomic RNA synthesis. *J.Virol.* **78**:980-994
257. **Zust, R., T. B. Miller, S. J. Goebel, V. Thiel, and P. S. Masters. 2008.** Genetic interactions between  
an essential 3' cis-acting RNA pseudoknot, replicase gene products, and the extreme 3' end of the  
mouse coronavirus genome. *J.Virol.* **82**:1214-1228



# Chapter 3

## **Site-directed mutagenesis of the nidovirus replicative endoribonuclease NendoU exerts pleiotropic effects on the arterivirus life cycle**

Clara C. Posthuma,  
Danny D. Nedialkova,  
Jessika C. Zevenhoven-Dobbe,  
Jeroen H. Blokhuis,  
Alexander E. Gorbalenya, and  
Eric J. Snijder

J. Virol. 2006. 80(4):1653-61  
Reprinted with permission.

## ABSTRACT

The highly conserved NendoU replicative domain of nidoviruses (arteriviruses, coronaviruses, and roniviruses) belongs to a small protein family whose cellular branch is prototyped by XendoU, a *Xenopus laevis* endoribonuclease involved in nucleolar RNA processing. Recently, sequence-specific in vitro endoribonuclease activity was demonstrated for the NendoU-containing nonstructural protein (nsp) 15 of several coronaviruses. To investigate the biological role of this novel enzymatic activity, we have characterized a comprehensive set of arterivirus NendoU mutants. Deleting parts of the NendoU domain from nsp11 of equine arteritis virus was lethal. Site-directed mutagenesis of conserved residues exerted pleiotropic effects. In a first-cycle analysis, replacement of two conserved Asp residues in the C-terminal part of NendoU rendered viral RNA synthesis and virus production undetectable. In contrast, mutagenesis of other conserved residues, including two putative catalytic His residues that are absolutely conserved in NendoU and cellular homologs, produced viable mutants displaying reduced plaque sizes (20 to 80% reduction) and reduced yields of infectious progeny of up to 5 log units. A more detailed analysis of these mutants revealed a moderate reduction in RNA synthesis, with subgenomic RNA synthesis consistently being more strongly affected than genome replication. Our data suggest that the arterivirus nsp11 is a multifunctional protein with a key role in viral RNA synthesis and additional functions in the viral life cycle that are as yet poorly defined.

abstract

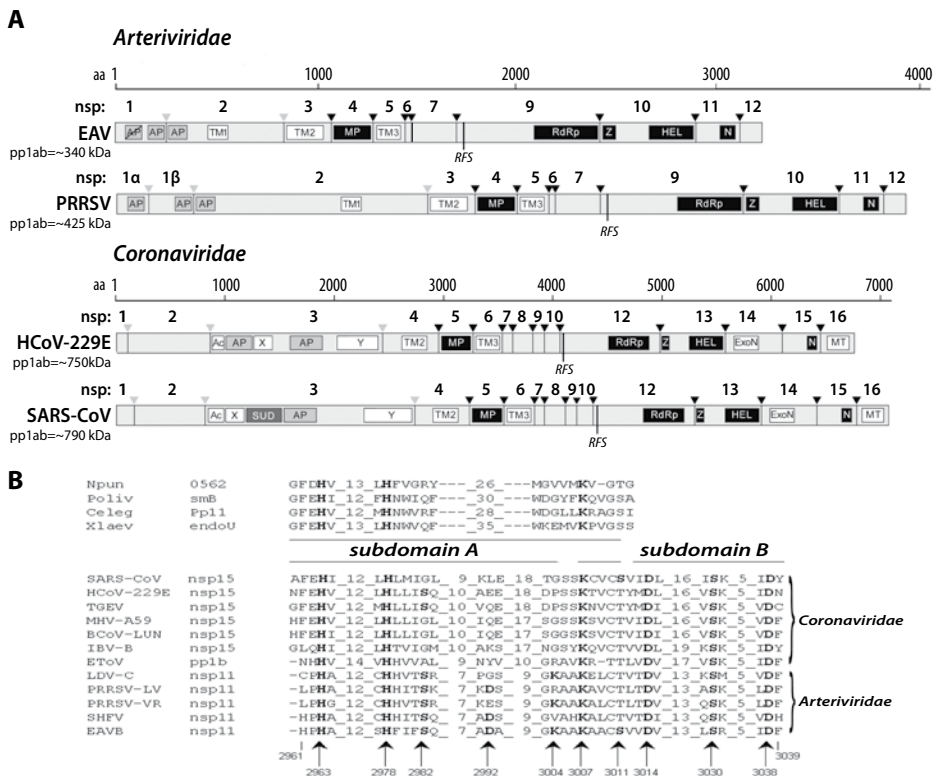
## INTRODUCTION

Nidoviruses are enveloped, positive-strand RNA viruses that have been grouped together on the basis of similarities in genome organization, the use of similar strategies for nonstructural and structural protein expression, and the presumed common ancestry of key replicative enzymes<sup>2,3,8,12,24</sup>. The order *Nidovirales* currently includes the families *Arteriviridae*, *Coronaviridae*, and *Roniviridae*. Notwithstanding their taxonomic unification, important differences between nidovirus subgroups exist at the levels of genome size, virion architecture, biology of virus-host interactions, and pathogenesis.

Nidovirus RNA synthesis in infected cells entails both amplification of the viral genome and the production of an extensive set of subgenomic (sg) mRNAs, presumably each from their specific negative-stranded intermediates<sup>17-19</sup>. RNA synthesis is driven by a set of nidovirus enzymes commonly referred to as “replicase”, which includes several activities that are rare or lacking in other RNA viruses<sup>23</sup>. The replicase is expressed from the incoming genome by translation of the large open reading frames (ORFs) 1a and 1b into polyproteins pp1a and pp1ab. The latter polypeptide is a C-terminally extended version of pp1a and is derived from a -1 ribosomal frameshift from ORF1a into overlapping ORF1b. The core of the nidovirus replicase polyprotein is formed by a conserved array of enzymatic domains consisting of (from N- to C-terminus) one or more papain-like “accessory proteinases”, a chymotrypsin-like “main proteinase”, the RNA-dependent RNA polymerase (RdRp), a helicase that contains a N-terminal zinc-binding domain, and an endoribonuclease called NendoU (for Nidovirus-specific endoribonuclease; see below)<sup>11,24,39</sup>. For nidovirus replicase polyproteins of over 6,500 amino acids, encoded by *Coronaviridae* and *Roniviridae* which have genomes of over 25 kilobases, several additional enzymatic activities have been predicted<sup>9,23,35,37</sup>. The fact that these enzymes have not been found in the proteome of the *Arteriviridae*, which have smaller genomes of between 12 and 16 kilobases, was linked to the postulated functional hierarchy of the predicted nidovirus “RNA-processing enzymes”, with the nidovirus-wide conserved domain (NendoU) playing a key role in the initiation of a cascade of reactions<sup>23</sup>.

The replicase polyproteins of equine arteritis virus (EAV), the arterivirus prototype, are autocatalytically cleaved into nonstructural proteins (nsps) 1 to 12 by three proteinases residing in nsp1, nsp2, and nsp4 (for a review, see reference<sup>40</sup>). The EAV structural proteins are encoded by seven ORFs in the 3' proximal quarter of the genome and are expressed from a nested set of sg mRNAs. These are not only 3' coterminal with the genome but also carry a common 211-nucleotide (nt) 5' leader sequence, which is identical to the 5' proximal region of the genome<sup>5</sup>. The sg mRNAs are produced via a unique mechanism that presumably involves discontinuous extension of minus strand RNA synthesis to produce sg-length minus strand templates for the synthesis of sg mRNAs<sup>17-19,22</sup>. The discontinuous step in minus-strand RNA synthesis, during which sequences that are noncontiguous in the antigenome are fused to give the sg-length minus strand RNAs, is guided by a base-pairing interaction between complementary transcription-regulating sequences (TRSs)





**Figure 1.** Nidovirus replicase comparison and sequence alignment of the EAV NendoU domain. (A) The replicase gene organization of the arteriviruses EAV and porcine reproductive and respiratory syndrome virus (PRRSV) and coronaviruses HCoV-229E and SARS-CoV are depicted in the form of the polyprotein pp1ab. The border between ORF1a- and ORF1b-encoded residues is indicated as RFS (ribosomal frameshift). Gray and black arrows represent the sites that are cleaved by accessory papain-like proteinases (AP) and the main ("3C-like") proteinases (MP), respectively. The proteolytic cleavage products are numbered, and within the cleavage products, the location of domains that have been identified as structurally or functionally related are indicated in black boxes. These include the four nidovirus-wide conserved domains encoded by ORF1b (RdRp, zinc-binding domain [Z], helicase [HEL], and NendoU [N]), putative transmembrane domains (TM), coronavirus-specific conserved domains (Ac, X, Y, ExoN, and MT), and one SARS-CoV specific insertion (SUD). For details, see references<sup>11,23,24,39</sup> and references therein. Note that the arterivirus and coronavirus polyproteins are drawn to different scales. (B) Sequence alignment of the most conserved regions of the NendoU domain of several nidoviruses with cellular proteins from the XendoU family. Subdomains A (present in all proteins) and B (nidovirus-specific) are indicated. Key conserved residues are shown in bold. In the EAV sequence (bottom line), arrows indicate the amino acids targeted in this study by site-directed mutagenesis; amino acid numbers refer to their position in EAV pp1ab. Abbreviations and accession numbers are as follows: Npun 0562, hypothetical protein of *Nostoc punctiforme*, ZP\_00106190; Poliv smB, pancreatic protein of *Paralichthys olivaceus*, BAA88246; Celeg Pp11, placental protein 11-like precursor of *Caenorhabditis elegans*, NP\_492590; Xlaev endoU, endoU protein of *Xenopus laevis*, CAD45344; SARS-CoV, NC\_004718; HCoV-229E, human coronavirus 229E, NC\_002645; TGEV, transmissible gastroenteritis virus, NC\_002306; MHV-A59, mouse hepatitis virus A59, NC\_001846; BCoV-LUN, bovine coronavirus LUN, AF391542; IBV-B, avian infectious bronchitis virus Beaudette, NC\_001451; EToV, equine torovirus, X52374; LDV-C, lactate dehydrogenase-elevating virus type C, NC\_002534; PRRSV-LV, porcine reproductive and respiratory syndrome virus, strain Lelystad, M96262; PRRSV-VR, porcine reproductive and respiratory syndrome virus, strain ATCC VR2332, AAD12125; SHFV, simian hemorrhagic fever virus, NC\_003092; and EAVB equine arteritis virus strain Bucyrus, NC\_002532.

present at the 3' end of the nascent minus strand (anti-body TRS) and in the leader region of the genomic template (leader TRS)<sup>17,27,32</sup>.

Four conserved domains, residing in proteins currently known as nsp9 to nsp11, were originally identified when the EAV ORF1b-encoded polypeptide was compared to its coronavirus and torovirus counterparts<sup>8</sup> (Fig. 1A): the RdRp in nsp9, a putative multinuclear zinc-finger and RNA helicase domain that are combined in nsp10<sup>20,21</sup>, and a nidovirus-wide conserved domain in the C-terminal part of nsp11 (between His-2961 and Phe-3039, using the pp1ab amino acid numbering). This latter domain has no known homologs in other groups of RNA viruses and thus it can be regarded as a genetic marker of the *Nidovirales*. A distant relationship between this nidovirus-specific domain and a family of cellular proteins prototyped by a *Xenopus laevis* endoribonuclease (XendoU)<sup>15</sup>, was recently identified<sup>23</sup>. XendoU is a Mn<sup>2+</sup>-dependent RNase that cleaves at U stretches and releases 2'-3'-cyclic phosphodiester products upon the processing of intron-encoded small nucleolar RNAs<sup>15</sup>. Its nidovirus homolog was dubbed "NendoU"<sup>13</sup> and using SARS-coronavirus and three other coronaviruses, two groups independently verified the *in vitro* endoribonuclease activity of the coronavirus nsp15 replicase subunit, which contains the NendoU domain<sup>1,13</sup>. Also, coronavirus NendoU was shown to be Mn<sup>2+</sup> dependent, produce molecules with 2'-3' cyclic phosphate ends, and cleave upstream and downstream of uridylates in GU or GUU sequences. Ivanov and colleagues<sup>13</sup> reported a preference of NendoU for double-stranded over single-stranded RNA molecules, but a slight preference for single-stranded RNA was claimed by Bhardwaj and colleagues<sup>1</sup>. Three residues (His-162, His-178, and Lys-224 in XendoU) that are absolutely conserved in proteins of the XendoU/NendoU family were implicated in catalysis<sup>10,23</sup>. Their replacement was not tolerated in coronavirus NendoU<sup>13</sup> and XendoU<sup>10</sup>. Using an infectious cDNA clone of human coronavirus 229E (HCoV-229E), a single substitution of a conserved NendoU Asp (Asp-6408; Fig. 1B), which also abolished endonuclease function in the *in vitro* assay, was found to completely block viral RNA synthesis and virus production<sup>13</sup>. This amino acid and a second conserved Asp residue (Asp-6435 in HCoV-229E; Fig. 1B) reside in the C-terminal part of the NendoU domain, which is well conserved in all nidoviruses, but cannot be confidently aligned with the corresponding region in the cellular homologs (see below).

To obtain more information on the role of NendoU in the life cycle of nidoviruses in general and of arteriviruses in particular, we performed an extensive site-directed mutagenesis study. Using our EAV reverse genetics system, the domain was partially or completely deleted and a variety of point mutations were introduced at the position of key conserved residues, including those presumably involved in catalysis. Mutants were tested for genome and sg RNA synthesis and, when viable, production of progeny virus. Whereas deletions and certain point mutations blocked all detectable viral RNA synthesis, other substitutions resulted in attenuated but viable virus mutants. A more detailed analysis of the latter mutants suggested a specific link between the NendoU domain and viral sg mRNA synthesis.

## MATERIALS AND METHODS

### Mutagenesis of nsp11 in the EAV full-length cDNA clone

Mutations in the EAV nsp11 NendoU domain were introduced in an appropriate shuttle vector by standard site-directed PCR mutagenesis as described by Landt *et al.*<sup>14</sup>. Restriction fragments containing the desired mutations were transferred to pEAV211, a wild-type (wt) EAV full-length cDNA<sup>26</sup>, after which all PCR-derived sequences were verified to exclude the introduction of unwanted additional nucleotide changes. The mutations engineered in the nsp11-coding region are listed in Table 1. pEAV211 was used as the wt control in all experiments. The NendoU deletion mutant was constructed by mutagenesis PCR, using a sense primer in which nucleotide 9091 was fused to nucleotide 9338 of the EAV genome, thereby making an in-frame deletion of the sequences encoding pp1ab residues His-2963 to Asp-3038. Similarly, mutants lacking the N-terminal part (His-2963 to Lys-3007) or the C-terminal part (Ser-3011 to Asp-3038) of the NendoU domain were constructed. The PCR products containing the deletions were transferred to pEAV211 using appropriate restriction sites.

### RNA transfection, EAV infection, immunoprecipitation and immunofluorescence analysis

Baby hamster kidney (BHK-21; ATCC CCL10) cells were electroporated with *in vitro* derived RNA transcripts of wt and mutant EAV full-length cDNA clones, as described previously<sup>29</sup>. Infection of BHK-21 cells with EAV-containing supernatants harvested after transfection, <sup>35</sup>S-labeling of protein synthesis, and immunoprecipitation analysis of transfected cells were performed by using the methods of de Vries and colleagues<sup>6</sup>. To ensure the quantitative nature of the immunoprecipitation, lysate and antiserum were titrated and it was confirmed that all of the target protein had been precipitated by performing a control reimmunoprecipitation on the supernatant of the primary immunoprecipitation. Immunofluorescence analyses with EAV-specific antisera against nsp3 (98E3) and N (Mab 3E2) were performed according to van der Meer and colleagues<sup>28</sup>. To monitor and compare transfection efficiencies, cells were labeled with anti-nsp3 serum 98E3 and Hoechst 33258 (1 µg/ml final concentration) to stain nuclei<sup>17</sup>. For plaque assays, subconfluent monolayers of BHK-21 cells were infected with 10-fold serial dilutions of wt or mutant EAV. Following a 1-h incubation, a 1% agar overlay was applied and cells were incubated at 39.5°C. Plaques were stained with 0.75% (w/v) crystal violet 4 days postinfection. To test nonviable mutants for a possible temperature-sensitive phenotype, plaque assays were also performed at 34°C.

## RNA isolation and analysis

Intracellular RNA was isolated by using the acidic phenol method as described by van Marle and colleagues<sup>32</sup>. RNA was separated in denaturing agarose-formaldehyde gels and detected by hybridization to a <sup>32</sup>P-labeled probe E154 (5'-TTGGTTCCTGGGTGGCTAATA-ACTACTT-3'), which is complementary to the 3' end of the EAV genome and recognizes the genome and all sg mRNAs. Phosphorimager screens were exposed to the hybridized gels and scanned with a Personal Molecular Imager FX (Bio-Rad). Band intensities were quantified with Quantity One v4.2.2 software (Bio-Rad).

The presence of the original NendoU mutation(s) in progeny virus was verified in the following way: supernatant harvested at the indicated time point was used to infect fresh BHK-21 cells. At 24 h post infection, total RNA was isolated and used in reverse transcription (RT)-PCR experiments with E533 (5'-CAATCACAGAAATAGCAAAATCAGC-3'; nt 9799-9823) as antisense primer for the RT and PCR and E519 (5'-CGAGGAAAAGTTT-GCCGCCGC-3'; nt 9082-9102) as sense primer for PCR. The PCR-products were purified and sequenced directly using standard protocols.

## RESULTS

### Deletion of the EAV NendoU domain is lethal

The boundaries of EAV nsp11 are Ser-2838 and Glu-3056 (using pp1ab numbering)<sup>30</sup>, with residues His-2961 to Phe-3039 forming the conserved 78-residue core of the NendoU domain. An alignment of the EAV NendoU domain with cellular and nidovirus homologs is shown in Fig. 1B. As explained in the introduction, the N-terminal part (His-2963 to Ser-3011; subdomain A) is conserved across the XendoU/NendoU family and contains the EAV equivalents (His-2963, His-2978, and Lys-3007) of the putative catalytic residues identified in XendoU (His-162, His-178, and Lys-224)<sup>10</sup> and HCoV-229E nsp15 (His-6345, His-6360, and Lys-6401)<sup>13</sup>. However, the C-terminal 30 residues (subdomain B) of the NendoU domain appear to be specific for nidoviruses and include two conserved Asp residues (Asp-3014 and Asp-3038), which were found to be equally indispensable for *in vitro* endoribonuclease activity<sup>13</sup>. Moreover, a reverse genetics experiment showed that an Asp-6408→Ala substitution completely blocked HCoV-229E RNA synthesis<sup>13</sup>.

As an initial test of its importance for EAV replication, in-frame deletion mutants of the EAV NendoU domain were engineered and transferred to the infectious cDNA clone<sup>29</sup>. Either the complete NendoU-coding sequence was deleted from the nsp11-coding region or the sequences roughly corresponding to subdomain A (deletion from His-2963 to Lys-3007) or subdomain B (Ser-3011 to Asp-3038) were deleted. Electroporation into BHK-21 cells of *in vitro*-transcribed RNA from these constructs and subsequent immunostaining showed that, in all three cases, the virus was no longer viable (data not shown), suggesting that the complete NendoU domain is essential for EAV replication.

**Table 1.** Overview of the genotype and first-cycle phenotype of EAV NendoU mutants used in this study<sup>a</sup>

Construct	Mutation	wt codon	Mutant codon	Nsp3 IFA <sup>b</sup>	N IFA <sup>b</sup>	Infectious progeny (PFU/ml) <sup>c</sup>	Summary of phenotype <sup>d</sup>
pEAV211	none (wild type)			+	+	+ (~10 <sup>7</sup> )	Wild type
H2963A	His-2963→Ala	CAU	GCA	+	+	+ (~10 <sup>2</sup> )	Small plaque size
H2963Q	His-2963→Gln	CAU	CAG	+	+	+ (~10 <sup>3</sup> )	Small plaque size
H2978A	His-2978→Ala	CAU	GCC	+	+	+ (~10 <sup>3</sup> )	Small plaque size
H2978Q	His-2978→Gln	CAU	CAA	+	+	+ (~10 <sup>2</sup> )	Small plaque size
H2963A/ H2978A	His-2963→Ala His-2978→Ala	CAU CAU	GCA GCC	+	+	+ (~10 <sup>2</sup> )	Small plaque size
S2982A	Ser-2982→Ala	UCC	GCC	+	+	+ (~10 <sup>7</sup> )	Wild type
D2992A	Asp-2992→Ala	GAC	GCA	-	-	-	Nonviable
D2992N	Asp-2992→Asn	GAC	AAC	+	+	+ (~10 <sup>7</sup> )	Wild type
D2992G	Asp-2992→Gly	GAC	GGC	+	+	+ (~10 <sup>7</sup> )	Wild type
K3004A	Lys-3004→Ala	AAA	GCA	+	+	+ (~10 <sup>6-107</sup> )	Wild type
K3004H	Lys-3004→His	AAA	CAU	+	+	+ (~10 <sup>6-107</sup> )	Wild type
K3004R	Lys-3004→Arg	AAA	AGA	+	+	+ (~10 <sup>7</sup> )	Wild type
K3007A	Lys-3007→Ala	AAA	GCC	+	+	+ (~10 <sup>2-103</sup> )	Small plaque size
K3007H	Lys-3007→His	AAA	CAU	+	+	+ (~10 <sup>2-103</sup> )	Small plaque size
K3007R	Lys-3007→Arg	AAA	CGG	+	+	+ (~10 <sup>3</sup> )	Small plaque size
S3011A	Ser-3011→Ala	AGC	GCC	+	+	+ (~10 <sup>5</sup> )	Intermediate plaque size
S3011T	Ser-3011→Thr	AGC	ACC	+	+	+ (~10 <sup>7</sup> )	Wild type
S3011C	Ser-3011→Cys	AGC	UGC	+	+	+ (~10 <sup>5</sup> )	Intermediate plaque size
D3014A	Asp-3014→Ala	GAU	GCU	-	-	-	Nonviable
D3014E	Asp-3014→Glu	GAU	GAG	-	-	-	Nonviable
D3014N	Asp-3014→Asn	GAU	AAC	-	-	-	Nonviable
S3030A	Ser-3030→Ala	AGU	GCC	+	+	+ (~10 <sup>6</sup> )	Intermediate plaque size
S3030T	Ser-3030→Thr	AGU	ACG	+	+	+ (~10 <sup>5</sup> )	Intermediate plaque size
S3030C	Ser-3030→Cys	AGU	UGC	+	+	+ (~10 <sup>6</sup> )	Intermediate plaque size
D3038A	Asp-3038→Ala	GAU	GCU	-	-	-	Nonviable
D3038E	Asp-3038→Glu	GAU	GAA	-	-	-	Nonviable
D3038N	Asp-3038→Asn	GAU	AAU	-	-	-	Nonviable

<sup>a</sup> Data from two independent transfection experiments are summarized. +, positive; -, negative.

<sup>b</sup> Transfected cells were analyzed after one cycle of replication at 19 h posttransfection. IFA, immunofluorescence assay.

<sup>c</sup> Results indicate the detection of infectious progeny (plaque assays) in medium from transfected cells harvested at 19 h posttransfection; approximate virus titers are shown between brackets.

<sup>d</sup> Nonviable indicates that no replication detected in immunofluorescence assays for nsp3 and N protein (see Fig. 4), and no infectious progeny were detected at 19 h post-transfection; small plaques were approximately 20% of the wt plaque size. The diameter of intermediate-size plaques ranged from 30 to 80% of the wt plaque size.



Formally, detrimental effects of these deletions (and other lethal mutations; see below) at the level of pp1ab proteolysis cannot be excluded even though their C-terminal border was relatively distant from the nsp11/12 cleavage site. Unfortunately, straightforward analysis in a *trans*-cleavage assay of the replicase processing event at the nsp11/12 junction was previously found to be impossible<sup>30</sup>.

On the basis of the alignment presented in Fig. 1B, both absolutely and less-conserved residues of EAV NendoU were targeted by site-directed mutagenesis and reverse genetics (Table 1). Each of the five invariant residues was replaced by Ala and one or two additional conservative substitutions. Also some less conserved and variable residues were replaced as a control for the alignment and to confirm that amino acid substitutions in general could be tolerated in this part of the polyprotein. Mutations were introduced into an EAV infectious cDNA clone and, after electroporation of *in vitro*-transcribed RNA, virus viability was analyzed at different time points after transfection. The initial screening was done on the basis of a double immunofluorescence assay for replicase subunit nsp3 and the nucleocapsid (N) protein, which can be used as markers for the synthesis of genome and sg mRNA<sup>7</sup>, respectively<sup>31</sup>. Cell culture supernatants were harvested around the end of the first replication cycle (19 h posttransfection) and tested in a plaque assay for the production of infectious virus progeny. Table 1 summarizes the mutations that were tested and the results of the first-cycle analysis that were obtained from two independent transfection experiments.

### **Both conserved Asp residues in NendoU subdomain B are critically involved in arterivirus RNA synthesis**

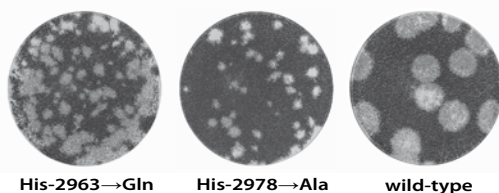
We first probed the importance of the pair of Asp residues (Asp-3014 and Asp-3038) in subdomain B that are fully conserved in the NendoU domain of all known nidoviruses. Both conservative (Asp→Asn or Asp→Glu) and non-conservative (Asp→Ala) substitutions were tested, but none of these were tolerated. In a first-cycle analysis, no sign of nsp3 or N protein synthesis was detected in immunofluorescence assays of transfected cells and the production of infectious progeny virus could not be demonstrated in plaque assays (Table 1). Ser-3030 of EAV NendoU is the third subdomain B residue that is fully conserved in all nidoviruses. Also, this amino acid was replaced with Ala and, more conservatively, with Thr and Cys. In all three cases, replication of the mutant virus was significantly reduced, with a plaque size that was 30-60% compared to that of the wt control (termed “intermediate size” in Table 1) and virus titers being 10- to 100-fold reduced. Another Ser residue that was probed, Ser-3011 of EAV NendoU, is located at the border between subdomains A and B. A Thr residue is found at this position in some cellular proteins of the XendoU family and in all nidovirus homologs, with the exception of SARS-CoV (Fig. 1B). To probe the importance of this residue and to corroborate the quality of the alignment in Fig. 1B, Ser-3011 was replaced by Thr. This indeed produced a mutant virus with a phenotype that could not be discriminated from that of the wt control, both in terms of virus titer that was obtained and in terms of plaque size. In contrast,

substitutions of the Ser-3011 residue by Ala or Cys were less well tolerated: replication of these mutant viruses was delayed, plaques were 60-80% of the wt size, and virus titers were ~100-fold reduced. Taken together, our data confirmed and extended the general importance for nidovirus RNA synthesis of NendoU subdomain B, and in particular its pair of conserved Asp residues.

### Substitution of putative catalytic residues in NendoU subdomain A is not lethal

Obviously, the EAV equivalents (His-2963, His-2978, and Lys-3007) of the three residues previously implicated in XendoU/NendoU catalysis (see above) were primary targets for site-directed mutagenesis of subdomain A. Both His residues were replaced with Ala or Gln, whereas Lys-3007 was substituted with Ala, His, or Arg. However, none of these seven substitutions was lethal, although all mutant viruses were severely impaired compared to the wt control. Irrespective of the replacement used, plaque sizes were reduced to approximately 20% of the wt plaque size and virus titers in the order of  $10^2$  to  $10^3$  PFU/ml were measured in transfected cell culture supernatant harvested at 19 h posttransfection, a reduction of 4 to 5 log compared to the wt virus. The reduced plaque size is illustrated in Fig. 2 for mutants His-2963→Gln and His-2978→Ala. There was no apparent synergistic effect of combining the Ala substitutions of both conserved His residues into a double mutant (mutant H2963A/H2978A): this virus displayed a similar small plaque size phenotype, and also produced  $\sim 10^2$  PFU/ml by 19 h posttransfection.

To verify the presence of the original mutations in the progeny of the mutants, fresh BHK-21 cells were infected with transfection supernatants harvested after 46 to 72 h. Intracellular RNA was isolated from these infected cells after 24 hours, and the direct sequence analysis of RT-PCR products derived from this material confirmed that reversion to the wt residue had not occurred for any of the mutants with small plaque sizes. In fact, the two His→Ala mutations in this double H2963A/H2978A mutant were found to be stable for up to 10 passages in BHK-21 cells (data not shown). In conclusion, although the substitution of equivalent residues in both XendoU and coronavirus NendoU reduced *in vitro* endoribonuclease activity to undetectable levels<sup>10,13</sup>, our *in vivo* data revealed that



**Figure 2.** Plaque morphology of selected EAV NendoU mutants. The plaque sizes from mutants carrying mutations at the position of the putative catalytic residues His-2963 and His-2978 were approximately fivefold smaller than those of the wt virus. Plaque sizes of mutants carrying other substitutions at the positions of His-2963 and His-2978 or replacements of Lys-3007 were comparable to the plaque sizes depicted in this figure.

EAV mutants carrying such mutations are capable of a consistently low but significant level of replication.

In addition to the replacement of the conserved subdomain A residues, several additional EAV NendoU residues were targeted, mainly as controls for the alignment and the substitutions described above. Replacement of Lys-3004, serving as a control for Lys-3007 mutagenesis, did not have any effect on virus viability, irrespective of whether Ala, His, or Asn was introduced. Also, substitution of a Ser residue (Ser-2982→Ala) that is fully conserved in *Arteriviridae* but not in *Coronaviridae* proved to be neutral with regard to virus replication. A final control substitution, replacement of the charged Asp-2992 by Ala, resulted in a nonviable virus. However, upon closer inspection of the alignment (Fig. 1B), we decided to also replace this residue with Gly (present in the mouse arterivirus lactate dehydrogenase-elevating virus) and Asn (which is more conservative compared to the Glu or Gln present at this position in most other nidoviruses). Both the Asp-2992→Gly and the Asp-2992→Asn mutant proved to have wt phenotype. Since, in contrast to Ala, both Gly and Asn frequently reside in  $\beta$ -turns, this result suggests that this property may be important for the residue at this position.

In summary, the EAV system was found to tolerate the substitution of three (putative) active site residues of NendoU which have been absolutely conserved throughout nidovirus evolution. Although replication of these mutants was profoundly impaired, as evident from reduced progeny titers and smaller plaque sizes, this phenotype was in clear contrast with that described above for the mutants carrying substitutions of the pair of conserved Asp residues in subdomain B.

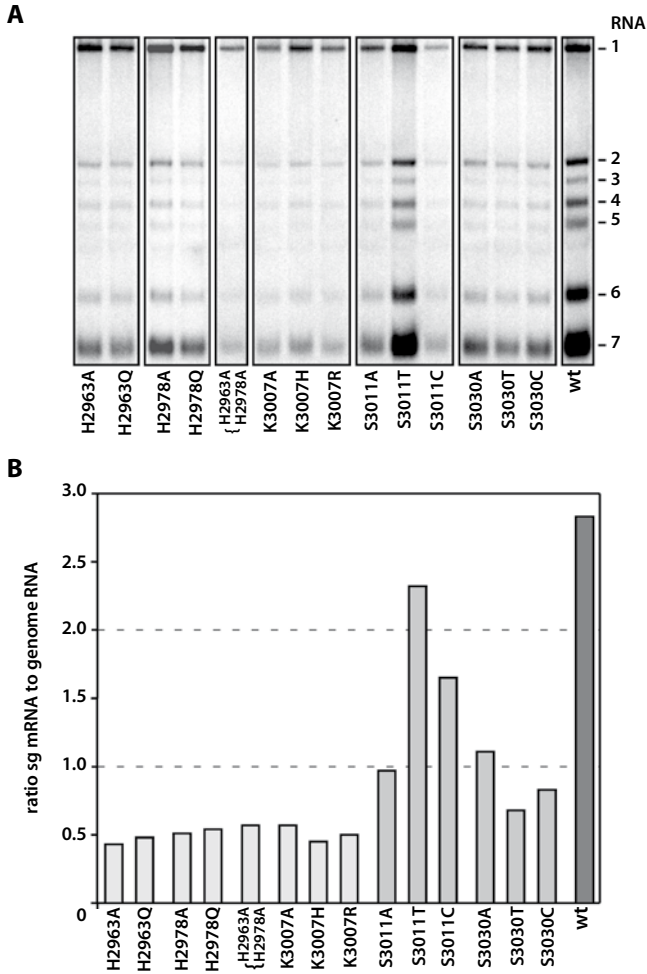
### **Selective reduction of subgenomic RNA synthesis in mutants with a reduced plaque size**

The intracellular RNA synthesis of selected mutants was analyzed by isolating total RNA from transfected BHK-21 cells at 13 h post electroporation, which, in electroporation experiments, approximately equals the peak of RNA synthesis during the first cycle of replication. The viral genome (RNA1) and sg mRNAs (RNAs 2 to 7) were separated in agarose gels and visualized by hybridization. No virus-specific RNAs were detected for the nonviable or quasi-viable Asp-3014 and Asp-3038 mutants (data not shown). In line with its previously described wt phenotype, an RNA pattern very similar to that of wt virus was observed for the Ser-3011→Thr mutant. All other mutants analyzed, including those with substitutions at the positions of His-2963, His-2978, and Lys-3007, showed a modest but significant reduction of viral RNA synthesis compared to the wt control (Fig 3A).

The relative amounts of genomic RNA and the three most abundant sg mRNAs (mRNAs 2, 6 and 7) were quantified by phosphorimager analysis. When the amount of genome RNA in each lane was set at 100%, the usual value of between 2:1 and 3:1 was obtained for the ratio of sg mRNAs versus genome RNA in the wt control and mutant S3011T, which has a wt phenotype (Fig. 3B). However, in all other mutants, a specific reduction of sg mRNA synthesis was detected. This effect was most prominent in the mutants that

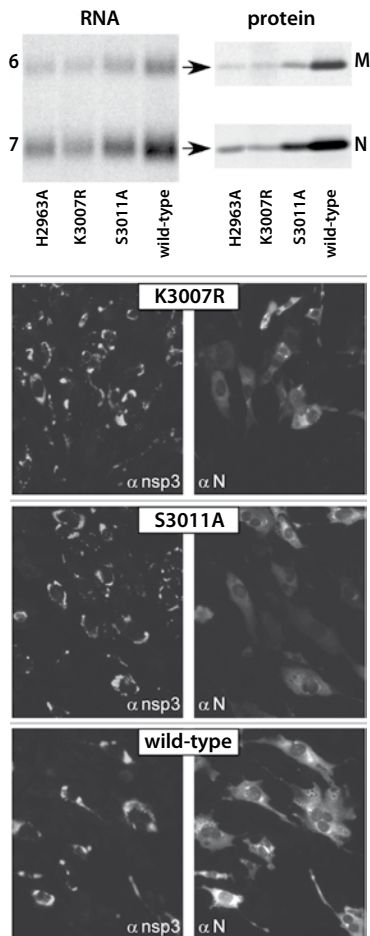
produced the smallest plaques, i.e. those with substitutions of His-2963, His-2978, Lys-3007, and the His-2963/His-2978 double mutant. The reduction of sg mRNA synthesis was less striking but still consistently observed for the Ser-3011 and Ser-3030 mutants, which had an intermediate plaque size (Table 1).

To make it possible to monitor viral RNA synthesis beyond the end of the first replication cycle, and without virus spreading to initially untransfected cells, we engineered



**Figure 3.** Analysis of intracellular RNA synthesis for selected EAV NendoU mutants. (A) Hybridization analysis of intracellular RNA isolated from transfected BHK-21 cells at 13 h post-transfection. The RNA was hybridized to an oligonucleotide complementary to the 3' end of the genome, recognizing both the genomic (RNA1) and sg mRNAs 2 to 7. (B) Reduced sg mRNA synthesis by nsp11 mutants with a small or intermediate plaque size phenotype. Using the gel depicted in Fig. 3A, the signals for the genomic RNA and the most abundant sg mRNAs (2, 6, and 7) in the same lane were quantified. The graph shows the ratio (RNA2 + RNA6 + RNA7)/RNA1 for each mutant and the wt control and illustrates the selective reduction of sg mRNA synthesis in NendoU mutants with a small or intermediate plaque size phenotype.

mutants that were negative for virus production. These constructs were based on the previously described ORF6 knock-out mutant ( $\Delta$ ORF6)<sup>16</sup>, which lacks membrane (M) protein expression and therefore is unable to produce progeny virus<sup>36,38</sup>. We transferred the NendoU His-2963→Ala/His-2978→Ala double mutation and the Lys-3007→Ala single mutation to the  $\Delta$ ORF6 backbone and compared RNA synthesis to that of the parental  $\Delta$ ORF6 mutant, which contains a wt replicase gene. Viral RNA synthesis was analyzed at 15, 18, 21, and 24 h post transfection and a selective reduction in sg mRNA synthesis was again observed for both NendoU mutants at all four time points (data not shown). This indicated that the earlier measurements (Fig. 3B) were not likely to be compromised by the possibility that sg RNA synthesis of these generally impaired NendoU mutants simply lagged behind. Compared to the wt control, these mutants consistently underproduced sg mRNAs, suggesting the direct or indirect involvement of NendoU in a process that is relatively more important for sg mRNA production than for genome RNA synthesis.



**Figure 4.** Analysis of structural protein expression of selected EAV NendoU mutants. The top panel shows a hybridization analysis of sg RNA6 and RNA7 isolated from transfected BHK-21 cells at 13 h post-transfection (also see Fig. 3). On the right side, the corresponding immunoprecipitation data are shown for the viral structural proteins M and N, which were <sup>35</sup>S-labeled from 9 to 13 h post-transfection. The bottom panel shows an immunofluorescence analysis (also at 13 h post-transfection) of cells transfected with small plaque size mutant Lys-3007→Arg, intermediate plaque size mutant Ser-3011→Ala, or the wild-type control. Cells were double-labeled for replicase subunit nsp3 and structural protein N. All images were recorded with the same exposure time and contrast/brightness settings. Whereas cells transfected with the wild-type control were double positive for replicase and N protein, many cells transfected with the mutants were fully positive for replicase but either still negative or only weakly positive for the N protein.

Structural protein synthesis was studied in more detail for a selection of mutants (Fig. 4). Immunofluorescence analysis confirmed that N protein expression of mutants with small and intermediate plaque sizes was reduced and lagged behind. Whereas all cells in the wild-type control were double-positive for replicase and N protein at 13 h post transfection, many cells transfected with the mutants were fully positive for replicase, but either still negative or only weakly positive for the N protein (Fig. 4). Only at 20-24 h posttransfection all cells had become clearly double-positive for both marker proteins (data not shown). In a combined RNA and immunoprecipitation analysis at 13 h posttransfection (Fig. 4) using samples from equal numbers of transfected cells, the amounts of sg RNA 6 and 7 accumulated by the small plaque mutants His-2963→Ala and Lys-3007→Arg and intermediate plaque size mutant Ser-3011→Ala was compared to the synthesis of the corresponding structural proteins M and N. Proteins were <sup>35</sup>S-labeled from 9 to 13 h post transfection and immunoprecipitated with monospecific rabbit antisera<sup>6</sup>. Compared to the wild-type control, the structural protein synthesis of the three mutants was clearly reduced but not to an extent that would readily explain the large differences (up to 5 log) in infectious progeny titers. When comparing mutants with small and intermediate plaque sizes (showing differences of 2 to 3 log units in infectious progeny titers), sg RNA and structural protein levels of the latter were found to be threefold higher at most. Finally, an anti-nsp11 serum was used to confirm that small and intermediate plaque size mutants produced similar amounts of cleaved nsp11 compared to the wild-type control (data not shown), making it unlikely that an nsp11 processing defect is the explanation for the phenotypes of these mutants.

## DISCUSSION

### Pleiotropic effects upon EAV NendoU mutagenesis

In the wake of the 2003 SARS outbreak, the increased interest in the giant and poorly characterized coronavirus replicase triggered an in-depth (re)analysis of the nonstructural proteins of SARS-CoV, other coronaviruses, and other nidoviruses. As a direct result, the domain previously referred to as “conserved C-terminal nidovirus-specific domain”<sup>8</sup> was predicted and proven to have endoribonuclease activity<sup>1,13,23</sup>. Using a HCoV-229E reverse genetics system, Ivanov and colleagues<sup>13</sup> showed that a single Asp→Ala mutation in the nidovirus-specific subdomain B of NendoU completely blocked viral RNA synthesis. Thus, they provided the first experimental evidence for the expected key role of the highly conserved, though nidovirus-specific, NendoU domain in the viral life cycle.

In the present study, using the reverse genetics system for the arterivirus EAV, we have considerably extended and refined the *in vivo* analysis of NendoU function. As previously concluded by Ivanov and colleagues, the results obtained with NendoU deletion mutants and certain subdomain B point mutants indicated that NendoU is critically involved in general nidovirus RNA synthesis. Although the replacement of one of three putative catalytic residues in subdomain A also exerted a major effect on general virus viability,

it was somewhat surprising that only a modest effect on viral genome and sg mRNA synthesis was observed for these mutants (Fig. 3), raising new questions on the role of the NendoU domain in the nidovirus life cycle.

Formally, we cannot exclude (fundamental) differences between the role of NendoU in coronaviruses and arteriviruses and the possibility that equivalent mutations may produce different phenotypes in different systems. Reverse genetics data on coronavirus NendoU mutants carrying replacements of the conserved His and Lys residues in subdomain A have not yet been described. Nevertheless, the result obtained with the single HCoV-229E Asp mutant that was tested<sup>13</sup> was identical to that obtained with the corresponding mutant in the EAV system. On the arterivirus side, it is unfortunate that we have not yet succeeded in establishing an *in vitro* NendoU assay for EAV. This is due to complications during protein expression in *E. coli*, which apparently tolerates the expression of large amounts of an Asp-3014→Ala mutant protein, but not of the wt EAV nsp11 (data not shown). Although the experimental proof that EAV nsp11 indeed possesses endoribonuclease activity is lacking, all residues reported to be important for the *in vitro* NendoU activity of its coronavirus ortholog are conserved in arteriviruses (Fig. 1B). Thus, we consider it unlikely that this activity would ultimately turn out to be lacking in arterivirus nsp11.

### **Is the function of NendoU uniform in the life cycle of different nidoviruses?**

It is perhaps remarkable that replacements of the most conserved residues in the XendoU/NendoU protein family fully blocked the *in vitro* endonuclease activity but were not lethal for EAV in the reverse genetics approach used in this study. Arguably, this apparent contradiction may reflect the higher sensitivity of the *in vivo* assay compared to the *in vitro* system as employed by Ivanov and colleagues<sup>13</sup>. In the latter assay, replacement of the pair of His residues and the Lys in subdomain A and the pair of Asp residues in subdomain B yielded equally inactive enzymes. However, it is highly unlikely that NendoU (or XendoU) mutants with very low, but different residual activities (e.g. 1% or less of the wt activity) can be discriminated using this type of experiment. In contrast, our biological assays, and in particular the quantification of infectious progeny virus produced by the mutants, allowed us to establish differences of up to 5 log (Table 1) and can thus be considered to be more sensitive. Consequently, one explanation for the phenotypic differences between the Asp mutants and the His/Lys mutants would be a higher residual endoribonuclease activity of the latter, which would suffice to support a certain reduced level of virus replication. This explanation would imply that the Asp residues in subdomain B, which are lacking in the cellular branch of the XendoU/NendoU family, play a more important role in NendoU function than the fully conserved His and Lys residues in subdomain A, which were proposed to have a catalytic role on the basis of experimental data and theoretical considerations relating to the reaction mechanisms of other ribonucleases<sup>10</sup>.

Despite the common ancestry of arterivirus and coronavirus replicases, the polyprotein of coronaviruses is considerably larger and in addition to NendoU harbors a series

of other putative RNA-processing activities with possibly interconnected functions<sup>23</sup>. It should also be noted, as observed previously by de Vries and colleagues<sup>7</sup>, that the NendoU domain is associated with a different N-terminal domain when coronavirus nsp15 and arterivirus nsp11 are compared. Within the currently defined NendoU domain, the role of subdomain B and its pair of conserved Asp residues is intriguing. Gioia and coworkers<sup>10</sup> suggested that the two NendoU Asp residues might be the equivalents of two XendoU Glu residues (Glu-161 and Glu-167) which were proposed to participate in catalysis on the basis of mutagenesis studies. However, these two amino acids more or less flank the first conserved His residue (His-162 in XendoU) and thus occupy a completely different position in the primary structure of the protein compared to the pair of conserved Asp residues in NendoU subdomain B. Clearly, structural information for both cellular and nidoviral representatives of this novel endoribonuclease family will be required to understand the domain composition and catalytic mechanism of the enzyme, and the specific role of subdomain B in the viral subgroup. At present, it should even be considered that the latter domain, although structurally adjacent to NendoU, could be functionally separated and fulfill a different role that might be the basis for some of the phenotypic differences observed in our collection of EAV mutants.

### **Dissecting the role of NendoU in the EAV life cycle**

The pleiotropic effects observed upon EAV NendoU mutagenesis and their analysis using the complex system of a replicating RNA virus in its host cell make it difficult to assess at how many levels NendoU may actually influence EAV replication. Clearly, specific deletions and mutations in NendoU can abolish all viral RNA synthesis, supporting a general and critical role in this process. Furthermore, we observed a stronger reduction of sg mRNA synthesis compared to genome synthesis in a subset of the mutants, but these findings could also be attributed to the former process being more sensitive to a general debilitating effect of NendoU mutagenesis on virus viability. Still, it appears that the His-2963/His-2978/Lys-3007 mutants form a separate group and that these mutations exert a consistently stronger effect on sg RNA synthesis than e.g. substitutions of Ser-3011 or Ser-3030 (Fig. 3B). Also in terms of the production of infectious virus progeny (virus titers and plaque size) these mutants appear to segregate into two clusters (Table 1).

When comparing the wt virus, the intermediate plaque size mutants (replacements of Ser-3011 and Ser-3030), and the small plaque size mutants (replacements of His-2963, His-2978, and Lys-3007), it appears that relatively small differences in RNA synthesis (certainly not more than 10-fold; Fig. 3) are translated into unexpectedly large differences in virus production (up to 5 log units) (Table 1). These findings could be interpreted to imply that NendoU has an additional function downstream of viral RNA synthesis, for example, in a pathway leading to virion assembly. However, it should also be noted that little is known about the process in which the seven EAV structural proteins, expressed from six sg mRNAs that are each made in constant but different quantities, interact to encapsidate the viral genome and form a virus particle<sup>36</sup>. In particular, it is unclear



whether the availability of any of these proteins is a rate-limiting factor for the assembly process. Therefore, it certainly cannot be excluded that a relatively small reduction in the synthesis of one or multiple sg mRNAs and a corresponding change in structural protein expression (Fig. 4) could actually cause a disproportional, much larger reduction in the production of infectious progeny.

In view of its conservation in and specificity for the order *Nidovirales*, the NendoU domain was proposed to be involved in a nidovirus-specific, conserved function. The synthesis of sg mRNAs, a hallmark of nidoviruses, is an example of such a function<sup>4,17-19,34</sup> and, based on the consensus sequence of the regulatory TRS motif, Ivanov and colleagues<sup>13</sup> proposed a role for the enzyme in the attenuation step during discontinuous sg minus-strand RNA synthesis. However, the EAV data presented in this study do not support a critical role for NendoU in sg RNA synthesis. Only a relatively small effect on RNA synthesis was observed for the putative active site mutants of EAV NendoU (Fig. 3B). Although, admittedly, sg RNA synthesis was affected to a greater extent than genome synthesis (see above), a much more pronounced effect would have been expected if NendoU were a key factor in nidovirus sg RNA production. Previous studies on other EAV replicase functions have shown that such specific effects can indeed be induced. For instance, deletion of the N-terminal replicase subunit nsp1 completely abolished EAV sg mRNA production without affecting genome replication<sup>25</sup>. Also, a Ser-2429→Pro mutation in the so-called “hinge region” of EAV nsp10, which connects the N-terminal (putative) zinc finger domain and the C-terminal helicase, resulted in an approximately 100-fold reduction of sg mRNA synthesis, while again genome replication was not affected<sup>29,33</sup>. Compared to these examples, the defects in RNA synthesis observed for the viable NendoU mutants described in this study should be classified as relatively minor and less specific for the process of sg mRNA synthesis.

In conclusion, notwithstanding the viability of many mutants, the EAV NendoU motif in nsp11 is an important replicase domain, although the consequences of replacement of some of its core residues were less black and white than expected beforehand. Since specific NendoU mutants were found to be replication competent and genetically stable, it will be interesting to analyze their long-term evolution both in BHK-21 cells and in more natural settings like equine cell lines, in which such mutants may have a more pronounced phenotype. In addition, identification of the natural substrates of coronavirus and arterivirus NendoU and structural information on members of the XendoU/NendoU family will be essential for a better understanding of the mutant phenotypes described in this study.

## ACKNOWLEDGEMENTS

We thank Cheng Kao, Konstantin Ivanov, John Ziebuhr, and Willy Spaan for helpful discussions and/or sharing data prior to publication. We also thank John Ziebuhr for critical reading of the manuscript and acknowledge the technical assistance of Erwin van den Born.

## REFERENCES

1. **Bhardwaj, K., L. A. Guarino, and C. C. Kao.** 2004. The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. *J.Virol.* **78**:12218-12224
2. **Cavanagh, D.** 1997. Nidovirales: a new order comprising Coronaviridae and Arteriviridae. *Arch. Virol.* **142**:629-633
3. **Cowley, J. A., C. M. Dimmock, K. M. Spann, and P. J. Walker.** 2000. Gill-associated virus of *Penaeus monodon* prawns: an invertebrate virus with ORF1a and ORF1b genes related to arteri- and coronaviruses. *J.Gen.Virol.* **81**:1473-1484
4. **Cowley, J. A., C. M. Dimmock, and P. J. Walker.** 2002. Gill-associated nidovirus of *Penaeus monodon* prawns transcribes 3'-coterminal subgenomic mRNAs that do not possess 5'-leader sequences. *J.Gen.Virol.* **83**:927-935
5. **de Vries, A. A., E. D. Chirnside, P. J. Bredenbeek, L. A. Gravestine, M. C. Horzinek, and W. J. Spaan.** 1990. All subgenomic mRNAs of equine arteritis virus contain a common leader sequence. *Nucleic Acids Res.* **18**:3241-3247
6. **de Vries, A. A., E. D. Chirnside, M. C. Horzinek, and P. J. Rottier.** 1992. Structural proteins of equine arteritis virus. *J.Virol.* **66**:6294-6303
7. **de Vries, A. A., M. C. Horzinek, P. J. M. Rottier, and R. J. de Groot.** 1997. The genome organization of the Nidovirales: similarities and differences between arteri-, toro-, and coronaviruses. *Semin.Virol.* **8**:33-47
8. **den Boon, J. A., E. J. Snijder, E. D. Chirnside, A. A. F. de Vries, M. C. Horzinek, and W. J. M. Spaan.** 1991. Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. *J.Virol.* **65**:2910-2920
9. **Feder, M., J. Pas, L. S. Wyrwicz, and J. M. Bujnicki.** 2003. Molecular phylogenetics of the RrmJ/fibrillarlin superfamily of ribose 2'-O-methyltransferases. *Gene* **302**:129-138
10. **Gioia, U., P. Laneve, M. Dlakic, M. Arceci, I. Bozzoni, and E. Caffarelli.** 2005. Functional characterization of XendoU, the endoribonuclease involved in small nucleolar RNA biosynthesis. *J.Biol. Chem.* **280**:18996-19002
11. **Gorbalenya, A. E.** 2001. Big nidovirus genome: When count and order of domains matter. *Adv. Exp.Biol.Med.* **494**:1-17
12. **Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov.** 1989. Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Res.* **17**:4847-4861
13. **Ivanov, K. A., T. Hertzog, M. Rozanov, S. Bayer, V. Thiel, A. E. Gorbalenya, and J. Ziebuhr.** 2004. Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc.Natl.Acad. Sci.U.S.A.* **101**:12694-12699
14. **Landt, O., H.-P. Grunert, and U. Hahn.** 1990. A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene* **96**:125-128
15. **Laneve, P., F. Altieri, M. E. Fiori, A. Scaloni, I. Bozzoni, and E. Caffarelli.** 2003. Purification, cloning, and characterization of XendoU, a novel endoribonuclease involved in processing of intron-encoded small nucleolar RNAs in *Xenopus laevis*. *J.Biol.Chem.* **278**:13026-13032
16. **Molenkamp, R., H. van Tol, B. C. D. Rozier, Y. van der Meer, W. J. M. Spaan, and E. J. Snijder.** 2000. The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription. *J.Gen.Virol.* **81**:2491-2496

17. **Pasternak, A. O., E. van den Born, W. J. M. Spaan, and E. J. Snijder.** 2001. Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. *EMBO J.* **20**:7220-7228
18. **Sawicki, S. G. and D. L. Sawicki.** 1995. Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. *Adv.Exp.Biol.Med.* **380**:499-506
19. **Sawicki, S. G. and D. L. Sawicki.** 2005. Coronavirus transcription: a perspective, p. 31-55. *In*: L. Enjuanes (ed.), *Coronavirus replication and reverse genetics*. Springer, Berlin.
20. **Seybert, A., C. C. Posthuma, L. C. van Dinten, E. J. Snijder, A. E. Gorbalenya, and J. Ziebuhr.** 2005. A complex zinc finger controls the enzymatic activities of nidovirus helicases. *J.Virol.* **79**:696-704
21. **Seybert, A., L. C. van Dinten, E. J. Snijder, and J. Ziebuhr.** 2000. Biochemical characterization of the equine arteritis virus helicase suggests a close functional relationship between arterivirus and coronavirus helicases. *J.Virol.* **74**:9586-9593
22. **Siddell, S. G., J. Ziebuhr, and E. J. Snijder.** 2005. Coronaviruses, toroviruses, and arteriviruses, p. 823-856. *In*: B. W. Mahy and V. ter Meulen (eds.), *Topley and Wilson's microbiology and microbial infections; Virology volume*. 10 ed. Hodder Arnold, London.
23. **Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. M. Poon, Y. Guan, M. Rozanov, W. J. M. Spaan, and A. E. Gorbalenya.** 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J.Mol. Biol.* **331**:991-1004
24. **Snijder, E. J., S. G. Siddell, and A. E. Gorbalenya.** 2005. The order Nidovirales, p. 390-404. *In*: B. W. Mahy and V. ter Meulen (eds.), *Topley and Wilson's microbiology and microbial infections; Virology volume*. 10 ed. Hodder Arnold, London.
25. **Tijms, M. A., L. C. van Dinten, A. E. Gorbalenya, and E. J. Snijder.** 2001. A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus. *Proc.Natl.Acad.Sci.U.S.A.* **98**:1889-1894
26. **van den Born, E., A. P. Gultyaev, and E. J. Snijder.** 2004. Secondary structure and function of the 5'-proximal region of the equine arteritis virus RNA genome. *RNA* **10**:424-437
27. **van den Born, E., C. C. Posthuma, A. P. Gultyaev, and E. J. Snijder.** 2005. Discontinuous subgenomic RNA synthesis in arteriviruses is guided by an RNA hairpin structure located in the genomic leader region. *J.Virol.* **79**:6312-6324
28. **van der Meer, Y., H. van Tol, J. Krijnse Locker, and E. J. Snijder.** 1998. ORF1a-encoded replicase subunits are involved in the membrane association of the arterivirus replication complex. *J.Virol.* **72**:6689-6698
29. **van Dinten, L. C., J. A. den Boon, A. L. M. Wassenaar, W. J. M. Spaan, and E. J. Snijder.** 1997. An infectious arterivirus cDNA clone: identification of a replicase point mutation which abolishes discontinuous mRNA transcription. *Proc.Natl.Acad.Sci.U.S.A.* **94**:991-996
30. **van Dinten, L. C., S. Rensen, W. J. M. Spaan, A. E. Gorbalenya, and E. J. Snijder.** 1999. Proteolytic processing of the open reading frame 1b-encoded part of arterivirus replicase is mediated by nsp4 serine protease and is essential for virus replication. *J.Virol.* **73**:2027-2037
31. **van Dinten, L. C., H. van Tol, A. E. Gorbalenya, and E. J. Snijder.** 2000. The predicted metal-binding region of the arterivirus helicase protein is involved in subgenomic mRNA synthesis, genome replication, and virion biogenesis. *J.Virol.* **74**:5213-5223
32. **van Marle, G., J. C. Dobbe, A. P. Gultyaev, W. Luytjes, W. J. M. Spaan, and E. J. Snijder.** 1999. Arterivirus discontinuous mRNA transcription is guided by base-pairing between sense and antisense transcription-regulating sequences. *Proc.Natl.Acad.Sci.U.S.A.* **96**:12056-12061



33. **van Marle, G., L. C. van Dinten, W. Luytjes, W. J. M. Spaan, and E. J. Snijder.** 1999. Characterization of an equine arteritis virus replicase mutant defective in subgenomic mRNA synthesis. *J.Virol.* **73**:5274-5281
34. **van Vliet, A. L., S. L. Smits, P. J. Rottier, and R. J. de Groot.** 2002. Discontinuous and non-discontinuous subgenomic RNA transcription in a nidovirus. *EMBO J.* **21**:6571-6580
35. **von Grotthuss, M., L. S. Wyrwicz, and L. Rychlewski.** 2003. mRNA cap-1 methyltransferase in the SARS genome. *Cell* **113**:701-702
36. **Wieringa, R., A. A. F. de Vries, J. van der Meulen, G. J. Godeke, J. J. M. Onderwater, H. van Tol, H. K. Koerten, A. M. Mommaas, E. J. Snijder, and P. J. M. Rottier.** 2004. Structural protein requirements in equine arteritis virus assembly. *J.Virol.* **78**:13019-13027
37. **Yan, L., M. Velikanov, P. Flook, W. J. Zheng, S. Szalma, and S. Kahn.** 2003. Assessment of putative protein targets derived from the SARS genome. *FEBS Lett.* **554**:257-263
38. **Zevenhoven-Dobbe, J. C., S. Greve, H. van Tol, W. J. M. Spaan, and E. J. Snijder.** 2004. Rescue of disabled infectious single-cycle (DISC) Equine arteritis virus by using complementing cell lines that express minor structural glycoproteins. *J.Gen.Virol.* **85**:3709-3714
39. **Ziebuhr, J.** 2004. Molecular biology of severe acute respiratory syndrome coronavirus. *Curr.Opin. Microbiol.* **7**:412-419
40. **Ziebuhr, J., E. J. Snijder, and A. E. Gorbalenya.** 2000. Virus-encoded proteinases and proteolytic processing in the *Nidovirales*. *J.Gen.Virol.* **81**:853-879

# Chapter 4

## **Biochemical characterization of arterivirus nonstructural protein 11 reveals the nidovirus-wide conservation of a replicative endoribonuclease**

Danny D. Nedialkova,  
Rachel Ulferts,  
Erwin van den Born,  
Chris Lauber,  
Alexander E. Gorbalenya,  
John Ziebuhr, and  
Eric J. Snijder

J. Virol. 2009. 83(11):5671-82  
Reprinted with permission

## ABSTRACT

Nidoviruses (arteriviruses, coronaviruses, and roniviruses) are a phylogenetically compact but diverse group of positive-strand RNA viruses that includes important human and animal pathogens. Nidovirus RNA synthesis is mediated by a cytoplasmic membrane-associated replication/transcription complex that includes up to 16 viral nonstructural proteins (nsps), which carry common enzymatic activities, like the viral RNA polymerase, but also unusual and poorly understood RNA-processing functions. Of these, a conserved endoribonuclease (NendoU) is a major genetic marker that is unique to nidoviruses. NendoU activity was previously verified *in vitro* for the coronavirus nsp15, but not for any of its distantly related orthologs from other nidovirus lineages, like the arterivirus nsp11. Here, we show that the bacterially expressed nsp11 proteins of two arteriviruses, equine arteritis virus and porcine respiratory and reproductive syndrome virus, possess pyrimidine-specific endoribonuclease activity. RNA cleavage was independent of divalent cations *in vitro* and was greatly reduced by replacement of residues previously implicated in catalysis. Comparative characterization of the NendoU activity in arteriviruses and severe acute respiratory syndrome coronavirus revealed common and distinct features of their substrate requirements and reaction mechanism. Our data provide the first biochemical evidence of endoribonuclease activity associated with arterivirus nsp11 and support the conclusion that this remarkable RNA-processing enzyme, whose substrate in the infected cell remains to be identified, distinguishes nidoviruses from all other RNA viruses.

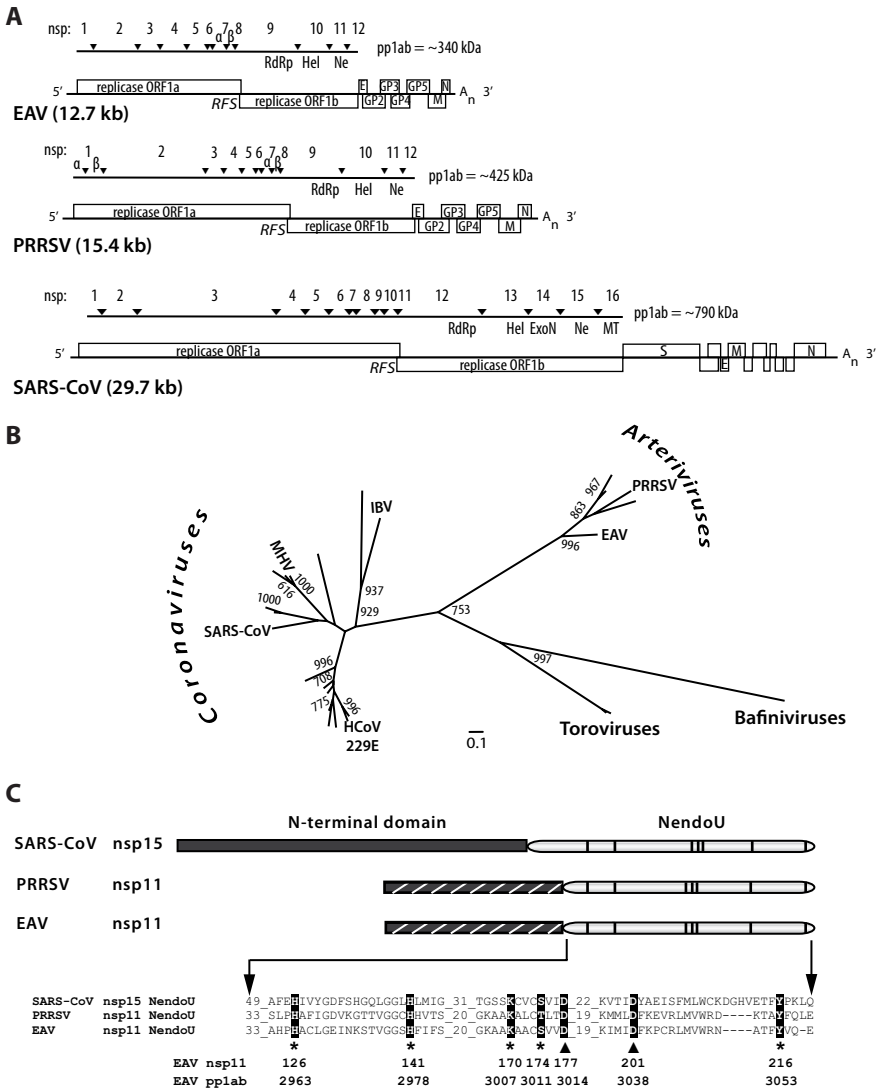
abstract

## INTRODUCTION

The replication of single-stranded RNA viruses with genomes of positive polarity (+RNA) relies on a process of RNA-templated RNA synthesis that takes place in the cytoplasm of the infected cell. Following translation of their genomes, +RNA viruses establish membrane-associated replication/transcription complexes that direct amplification of the viral genome and, in several cases, transcription of additional viral mRNAs. The key replication/transcription complex component in these processes is a virus-encoded RNA-dependent RNA polymerase (RdRp), but in addition, +RNA viruses generally encode multiple protein factors that support RdRp activity and/or are involved in the spatial and temporal regulation of viral RNA synthesis. Many of these enzymes were initially identified by comparative sequence analysis, subsequently supported by biochemical and molecular biological experiments. In nidoviruses, arguably the genetically most complex +RNA viruses<sup>12</sup>, examples of such accessory enzymes are a second, low fidelity RdRp<sup>15</sup>, as well as an unusual set of "RNA-processing enzymes"<sup>33</sup>, which are all genetically segregated in the viral replicase gene and were implicated in various steps of the nidovirus life cycle<sup>8,16,21,24</sup>.

The order *Nidovirales* comprises the distantly related families *Arteriviridae*, *Coronaviridae* (genera *Coronavirus* and *Torovirus*) and *Roniviridae* and includes the virus groups with the largest RNA genomes known to date<sup>12</sup>. The host range of *Nidovirales* extends from crustaceans to mammals, and the order includes pathogens associated with human respiratory infections, with diseases that can vary from common cold to fatal pneumonia. Despite major differences in virion size and morphology, as well as pathogenesis, the evolutionary relationship among nidoviruses is evident from the organization and composition of the large replicase gene, which consists of the two 5'-most open reading frames (ORFs) – ORF1a and ORF1b, of the polycistronic genome (Fig. 1A). Translation of the 5'-capped and 3'-polyadenylated nidovirus genomic RNA results in two large precursor polyproteins (pp1a and pp1ab) that are cleaved into 13 to 16 non-structural proteins (nsps) by virus-encoded proteases. The relative abundance of nidovirus nsps is intricately regulated by the proteolytic processing of the pp1a and pp1ab precursors (references<sup>35,39</sup> and references therein), as well as the efficiency of the -1 ribosomal frameshift required for the translation of ORF1b<sup>6</sup>. The nsps direct the formation of a membrane-anchored replication/transcription complex that mediates viral genome replication, as well as the synthesis of a 3'-coterminal nested set of subgenomic (sg) mRNAs<sup>23,31</sup>, from which the viral genes located downstream of the replicase gene are expressed.

The core enzymatic activities commonly required for viral RNA synthesis, such as the viral RdRp and helicase functions, are encoded in ORF1b of the nidovirus genome, together with an additional conserved domain that has no counterpart in the RNA virus world and is thus considered a genetic marker of the nidovirus order<sup>12,16</sup>. Bioinformatics revealed a distant relationship between this domain and a small family of prokaryotic and eukaryotic proteins, of which XendoU, an endoribonuclease derived from *Xenopus laevis* that plays a role in the production of intron-encoded small nucleolar RNAs<sup>19</sup>, is the prototype.



**Figure 1.** Conservation of the NendoU endoribonuclease across the order *Nidovirales*. (A) Genome organization of the arteriviruses EAV and PRRSV, as well as SARS-CoV. The ORFs, as well as the ORF1a/1b ribosomal frameshift site (RFS) and the 3' poly(A) tail (A<sub>n</sub>), are indicated. The ORFs encoding the viral structural proteins envelope (E), membrane (M), nucleocapsid (N), and glycoprotein 2 (GP<sub>2</sub>) to GP<sub>5</sub> for EAV and PRRSV and E, M, N, and spike (S) for SARS-CoV are designated. The organization of the respective replicase gene products is depicted in the form of pp1ab. The arrowheads represent sites cleaved by virus-encoded proteases, and the resulting nsps are numbered. The locations of the viral RdRp, helicase (Hel), ExoN, NendoU (Ne), and ribose-2'-O-methyltransferase (MT) domains (see the text) are indicated. Note that the SARS-CoV genome is drawn to a different scale than the EAV and PRRSV genomes. (B) Phylogenetic tree of the NendoU domains of arteriviruses, coronaviruses, toroviruses, and bafiniviruses. For clarity, species names are given only for coronaviruses, for which NendoU activity was previously documented (HCoV 229E, human coronavirus 229E; SARS-CoV; MHV; IBV, infectious bronchitis virus), as well as the arteriviruses whose NendoU activities were analyzed in this study (EAV and PRRSV). The sequence alignment used for construction of the tree, as well as the complete version of the tree, is included as supplemental material



(see Fig. S1 and S2 in the supplemental material). The numbers at branching points indicate the support of at least 600 out of 1,000 bootstraps, and the branch lengths of the tree represent the average numbers of substitutions per residue. (C) Domain organization of SARS-CoV nsp15, PRRSV nsp11, and EAV nsp11. Note the difference in size between the respective N-terminal domains. The homologous EAV and PRRSV nsp11 N-terminal domains (A.E. Gorbalenya, unpublished) are hatched. The partial sequence alignment shows the most conserved regions in the NendoU domains of the three viral proteins used in this study and was extracted from the alignment presented in Fig. S1 in the supplemental material. For EAV, the amino acid numbers given reflect the positions in EAV nsp11 (top row; used in this study) or in replicase polyprotein pp1ab (bottom row; used in previous studies). Conserved residues mutated in this study are highlighted. The asterisks indicate residues in EAV nsp11 NendoU that were replaced with Ala. The asterisks indicate proposed active-site residues, and nidovirus-wide conserved aspartic acid residues are indicated with triangles.

XendoU cleaves U stretches in the RNA substrate *in vitro* in a  $Mn^{2+}$ -dependent manner, releasing 2'-3'-cyclic phosphodiester products<sup>19</sup>, and a similar endoribonuclease activity was recently shown for a mammalian homolog of XendoU, human placental protein 11 (PP11)<sup>20</sup>. In coronaviruses, the proposed endoribonuclease activity maps to the C-terminal domain of nsp15, and uridylylate-specific cleavage of RNA was demonstrated *in vitro* for several nsp15 orthologs<sup>3,16,38</sup>. The enzyme was termed NendoU (*nidoviral endoribonuclease specific for U*), and its crucial role in the coronavirus life cycle was established using recombinant coronaviruses that expressed mutant forms of nsp15<sup>16,18</sup>. Biochemical studies have provided some insights into the catalytic activity of coronavirus NendoU's, such as their ability to process both single-stranded RNA (SSRNA) and double-stranded RNA (dsRNA) 3' of uridylylates and to release 2'-3'-cyclic phosphodiester products<sup>5,16</sup>. The crystal structures of nsp15 from mouse hepatitis virus (MHV) and severe acute respiratory syndrome coronavirus (SARS-CoV) were recently solved<sup>29,38</sup>, revealing an active site with two histidines and a lysine as likely catalytic residues. The spatial arrangement of the putative catalytic residues was strikingly similar to that found in RNase A, a feature also proposed for the catalytic site of XendoU<sup>28</sup>. The critical roles of these three highly conserved residues (see Fig. S1 in the supplemental material) in catalysis was supported by the results of site-directed mutagenesis studies<sup>11,16,38</sup>, and the importance of the two His residues for optimal viral RNA synthesis during MHV infection was established<sup>18</sup>. Coronavirus nsp15 proteins form hexamers via intermolecular interactions that are mediated by their N-terminal domains<sup>13,29,38</sup>. Hexamerization has been proposed to be required for optimal enzyme activity, most probably by stabilizing the catalytic site in its active conformation<sup>17</sup> and promoting substrate binding<sup>38</sup>, and a recent study identified numerous residues of SARS-CoV nsp15, mapping both to the N-terminus and the catalytic domain of the protein, as contributing to RNA binding by the hexamer<sup>4</sup>. Lastly, the presence of  $Mn^{2+}$  was shown to enhance coronavirus NendoU activity *in vitro*<sup>16,38</sup>, but its precise role in RNA cleavage remains unclear, since it appears to be dispensable for catalysis<sup>38</sup> and no metal-binding sites have been identified in any of the coronavirus enzymes characterized to date<sup>29,38</sup>.

The coronavirus NendoU domain has poorly characterized orthologs in other nidovirus lineages, like arteriviruses, toroviruses and white bream virus, the prototype bafinivirus<sup>32</sup>, which are all separated by large evolutionary distances (Fig. 1B; see Fig. S2 in the supple-

mental material). For example, the NendoU domains of coronavirus nsp15 and arterivirus nsp11 share only 26-27% of identical residues and are fused to a family-specific N-terminal domain (Fig. 1C). Clearly, the biochemical characterization of arterivirus nsp11 and its counterparts in other nidoviruses is a prerequisite to understanding the functional consequences of the observed sequence variation. The prior analysis of arterivirus nsp11 was limited to a reverse genetics study of NendoU mutants using the arterivirus prototype, equine arteritis virus (EAV)<sup>24</sup>, which revealed a key role in the EAV life cycle for the complete nsp11 subunit, as well as the specific importance of the NendoU domain and its catalytic residues for viral RNA synthesis. Characterization of the putative EAV nsp11 endoribonuclease activity *in vitro*, however, was previously hampered by the fact that recombinant nsp11 is extremely toxic to a variety of hosts (C.C. Posthuma, E. van den Born, and E. J. Snijder, unpublished observations). The use of a tightly regulated bacterial expression system has now enabled us to express and purify recombinant nsp11 for EAV and porcine reproductive and respiratory syndrome virus (PRRSV), two distantly related arteriviruses whose NendoU domains share ~52% amino acid identity. This technical advance allowed us to initiate the protein's functional characterization and here we provide the first direct evidence for the *in vitro* endoribonuclease activity of arterivirus nsp11. Furthermore, we have studied the enzyme's biochemical requirements for optimal RNA processing and the contribution of proposed active site residues to catalysis. Our comparative approach allowed us to identify biochemical properties common to the arteri- and coronavirus endoribonucleases, as well as distinct characteristics of the arterivirus NendoU, suggesting that it is able to process RNA efficiently in the absence of divalent cations by using a cleavage reaction mechanism resembling that of RNase A.

## MATERIALS AND METHODS

### Bioinformatics analyses

Complete genome sequences of corona-, toro-, bafini- and arteriviruses available on 16 April 2008 at the National Center for Biotechnology Information GenBank/RefSeq databases were downloaded into the Vialis platform (A.E.Gorbalenya, unpublished). An amino acid alignment of the NendoU domain was produced using the MUSCLE program<sup>9</sup> and alignment of poorly conserved columns was further manually refined to produce a final alignment (174 positions). A selection of 25 sequences representing virus species diversity (C. Lauber and A.E. Gorbalenya, unpublished) was used for phylogenetic analyses. The following genome sequences were used: bovine torovirus Breda-1 (NC\_007447); equine torovirus Berne (X52374); white bream virus (NC\_008516); infectious bronchitis virus strain Beaudette (NC\_001451); Beluga Whale coronavirus SW1 (EU111742); human coronavirus NL63 (NC\_005831.2); human coronavirus 229E (NC\_002645); *Miniopterus* bat coronavirus 1B (NC\_010436); *Miniopterus* bat coronavirus HKU8 (NC\_010438); *Rhinolophus* bat coronavirus HKU2 (NC\_009988); *Scotophilus* bat coronavirus 512/05 (DQ648858);

porcine epidemic diarrhoea coronavirus (NC\_003436); feline coronavirus (NC\_007025); mouse hepatitis virus A59 (NC\_001846); human coronavirus HKU1 (NC\_006577.2); human coronavirus OC43 (NC\_005147); *Tylonycteris* bat coronavirus HKU4 (EF065505); *Pipistrellus* bat coronavirus HKU5 (EF065509); *Rousettus* bat coronavirus HKU9 (EF065513); severe acute respiratory syndrome coronavirus (AY291315); porcine respiratory and reproductive syndrome virus, strain VR-2332 (AY150564); porcine respiratory and reproductive syndrome virus, strain Lelystad (AY588319); lactate dehydrogenase-elevating virus neuro-virulent type C (L13298); simian hemorrhagic fever virus (NC\_003092) and equine arteritis virus (NC\_002532.2). Utilizing the PhyML program<sup>14</sup>, a maximum likelihood phylogenetic tree was compiled using the WAG model of amino acid substitution<sup>37</sup> and rate heterogeneity among sites (gamma distribution with 8 categories). Support for internal nodes was measured by bootstrap analysis (1,000 replicates).

### Plasmid construction

Nucleotides 8735 to 9391 of the wild-type EAV infectious cDNA clone pEAV211<sup>36</sup>, encoding EAV nsp11, were PCR-amplified with oligonucleotides E780 (5'-CATGCCATGGGCCATCAC-CATCACCATCACTCCAACAAAATTCGTGCCTCCCG-3'), introducing a start codon followed by sequences specifying a (His)<sub>6</sub> tag, and E584 (5'-CCGGAATTCGCATGCTAGCTCACTCTTG-GACATAAAAGGTCGCG-3'), providing a stop codon. The fragments were cloned between the NcoI and EcoRI restriction sites of Gateway entry vector pENTR11 (Invitrogen, Carlsbad, CA, USA), and, after sequence verification, placed downstream of the T7 promoter in Gateway expression vector pDEST14. Translation of the resulting construct, pDEST14-EAV nsp11HN, produced a protein corresponding to amino acids Ser-2838 to Glu-3056 of replicase pp1ab of the EAV-Bucyrus strain (NCBI genome database accession number NC\_002532), preceded by an N-terminal Met-Gly-His<sub>6</sub> tag. Mutant pEAV211 plasmids encoding alanine substitutions in the EAV nsp11 region<sup>24</sup>(with the exception of Tyr-216 Ala) were used as templates for the generation of expression constructs encoding mutant derivatives of EAV nsp11. The Tyr-216 Ala substitution was newly engineered by using primers E780 and E1020 (5'-GGAATTCTCACTCTTGGACAGCAAAGGTCG-3'). The PRRSV nsp11-coding sequence (strain VR-2332; accession number AY150564) was PCR-amplified from a full-length cDNA clone<sup>22</sup> using oligonucleotides HN-fwd (5'-GGGGACAAGTTG-TACAAAAAGCAGGCTTCGAAGGAGATGCCACCATGAAACATCACCATCACCATCACGGGTC-GAGCTCTCCGCTCCC-3') and HN-rev (5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCT-TATTATTCAAGTTGAAATAGGCTGTTTTG-3'). The corresponding PCR fragment was recombined into Gateway entry vector pDONR201, the entry clone was sequenced, and the fragment was transferred to pDEST14. Nsp11 expressed from the resulting construct, pDEST14-PRRSV nsp11HN, corresponded to amino acids Gly-3585 to Glu-3807 of PRRSV-VR2332 replicase pp1ab, fused to an N-terminal Met-Lys-His<sub>6</sub> tag. The pDEST14 vector encoding wild-type nsp15 from SARS-CoV (Frankfurt-1 strain) was described previously<sup>29</sup>.

## Protein expression and purification

*Escherichia coli* BL21-AI cells (Invitrogen) were transformed with pDEST14-EAV nsp11HN, pDEST14-PRRSV nsp11HN and pDEST14-SARS-CoV nsp15HN, respectively, and plated on LB-agar containing 100 µg/ml ampicillin and 0.2% D-glucose. Subsequently, liquid cultures were inoculated by scraping colonies of freshly transformed cells into LB supplemented with 100 µg/ml of ampicillin and 0.2% D-glucose and grown at 37°C until the OD<sub>600</sub> equaled 0.6. The growth medium was exchanged with LB containing 100 µg/ml of ampicillin, the culture temperature was lowered to 16°C and protein expression was induced by addition of 0.2% L-arabinose (Sigma-Aldrich, St. Louis, MO, USA) for 5 h. Cells were then harvested by centrifugation and lysed in 20 mM Tris-HCl pH 8.0; 500 mM NaCl; 5% glycerol; 12.5 mM imidazole; 0.1% Triton-X100, supplemented with 50 U/ml DNase I (Invitrogen) and 1 mg/ml lysozyme (Sigma-Aldrich) for 1 h at 4°C. The Triton-X100 concentration was adjusted to 0.5%, lysates were incubated at 4°C for additional 30 min and subsequently clarified by centrifugation at 20,000g for 45 min. Recombinant proteins were purified from supernatants using Talon metal affinity resin (Clontech, Mountain View, CA, USA). The resin was washed extensively with 20 mM Tris-HCl pH 8.0; 500 mM NaCl; 5% glycerol; 12.5 mM imidazole, and the proteins were eluted in 20 mM HEPES, pH=7.4, 100 mM NaCl, 150 mM imidazole. The protein yield and purity of the preparations were assessed by SDS-PAGE and Coomassie Blue staining using bovine serum albumin (Bio-Rad, Hercules, CA, USA) as a concentration standard. The purified proteins were dialyzed against buffer containing 20 mM HEPES, pH 7.4; 100 mM NaCl; 50% glycerol, and stored at -20°C.

## RNA and DNA substrates

RNA and DNA oligonucleotides were chemically synthesized by Eurogentec (Seraing, Belgium). Substrate RNA2 (5'-CGCAGUGAGCUCCUAAUCGCCC-3') and markers RNAm2 (5'-CGCAGUU-3'), RNAm3 (5'-CGCAGU-3') and RNAm4 (5'-CGCAG-3') were described previously<sup>16</sup>. dsRNA substrates were generated by annealing labeled RNA2 and unlabeled RNA13 (5'-GGGCGAUUAGGAGCUCACUGCG-3') at equimolar concentrations. Oligoribonucleotides RNAsGUA (5'-AAAAAAGUAAAAA-3') and RNAsGCA (5'-AAAAAAGCAAAAA-3') were used to assess the preference of the enzymes for cleavage after uridylate or cytidylate. DNA oligonucleotides MHV324 (5'-GATCCAATTGTCGACGCGTG-3') and MHV325 (5'-AATTCACGCGTCGACAATTG-3'), which are partially complementary, were used to determine the activity of arterivirus nsp11 on ssDNA and dsDNA templates. RNA and DNA oligonucleotides were 5' end labeled using T4 polynucleotide kinase (PNK) (Invitrogen) and [ $\gamma$ -<sup>32</sup>P]ATP (PerkinElmer, Waltham, MA, USA) and were purified using RNase-free Micro Bio-spin 30 columns (Bio-Rad) prior to being used as substrates.

## Endoribonuclease assays

Proteins were incubated with the indicated concentration of 5'-<sup>32</sup>P-labeled substrate in 25 mM HEPES, pH 7.4; 50 mM NaCl, 1 mM DTT, 0.2 mM EDTA at 30°C. When MnCl<sub>2</sub> was present, EDTA was omitted from the reaction buffer. EAV nsp11 and SARS-CoV nsp15 were used at a final concentration of 0.3 μM, and PRRSV at 0.8 μM, unless otherwise indicated. Cleavage assays with SARS-CoV nsp15 were performed in the presence of 5 mM MnCl<sub>2</sub>. The reactions were stopped by the addition of an equal volume of Gel Loading Buffer II (95% formamide, 18 mM EDTA, 0.025% SDS; Ambion, Austin, TX, USA) and incubation for 3 min at 95°C. The reaction products were separated by electrophoresis in 20% polyacrylamide/7 M urea/1x TBE gels, supplemented with 0.1% SDS for the analysis of double-stranded substrate processing. The gels were exposed to phosphorimager screens, which were subsequently scanned using a Typhoon Variable Mode Imager (GE Healthcare, Chalfont St Giles, UK). Images were analyzed and quantified using ImageQuant TL software (GE Healthcare).

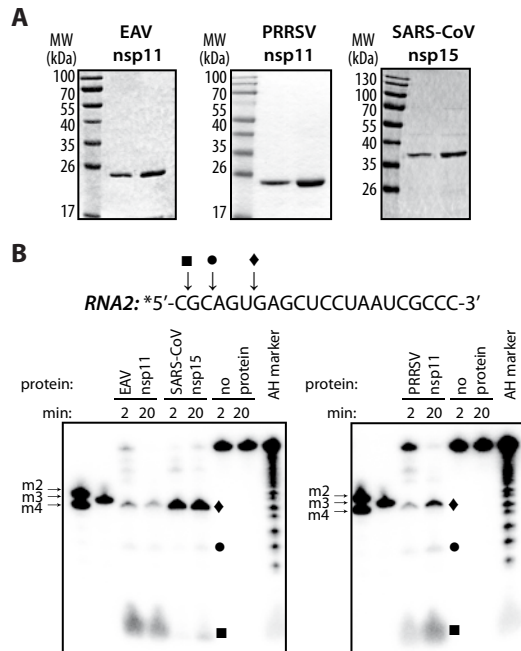
## Analysis of 3' termini of cleavage products

The analysis of the nature of the 3' ends of 5' cleavage products derived from NendoU-mediated RNA processing was performed essentially as previously described<sup>16</sup>. Unlabeled RNAsGUA substrate was incubated with recombinant EAV nsp11 or SARS-CoV nsp15. The reaction was stopped by heat denaturation for 4 min at 95°C and the cleavage products were incubated with T4 PNK buffer alone (70 mM Tris-HCl, pH 7.6; 10 mM MgCl<sub>2</sub>; 100 mM KCl; 1 mM 2-mercaptoethanol), T4 PNK buffer in the presence of 5 U of T4 PNK, or calf intestine alkaline phosphatase (CIAP) buffer (10 mM Tris-HCl, pH 7.5; 10 mM MgCl<sub>2</sub>) in the presence of 0.5 U CIAP (Fermentas, Burlington, ON, Canada), for 30 min at 37°C. The phosphatases were subsequently heat-inactivated at 95°C for 5 min. The RNAs were phenol-extracted, ethanol-precipitated using 10 μg glycogen (Roche, Basel, Switzerland) as a carrier, and ligated to [5'-<sup>32</sup>P]pCp using T4 RNA ligase (Fermentas) as previously described<sup>16</sup>. The reaction products were then resolved in denaturing polyacrylamide gels and visualized as described above.

## RESULTS

### Arterivirus nsp11 has Mn<sup>2+</sup>-independent endoribonuclease activity

Previous attempts to express and purify recombinant EAV nsp11 had been unsuccessful due to the high toxicity that the wild-type protein displays in prokaryotic and eukaryotic cells (Posthuma, van den Born, unpublished). We have now used the BL21-AI strain of *E. coli* as host for the T7 RNA polymerase-driven expression of N-terminally His<sub>6</sub>-tagged nsp11. The regulation of T7 polymerase expression in this BL21-derived strain is controlled by the



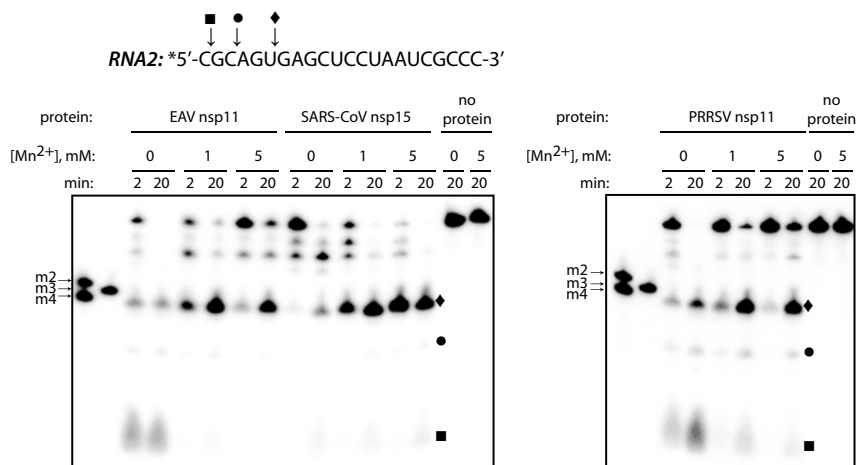
**Figure 2.** Purification and endoribonuclease activity of arterivirus nsp11. (A) Purified recombinant EAV nsp11 (0.5  $\mu$ g and 1  $\mu$ g), PRRSV nsp11 (2  $\mu$ g and 4  $\mu$ g), and SARS-CoV nsp15 (0.5  $\mu$ g and 1  $\mu$ g), each carrying an N-terminal His<sub>6</sub> tag, were analyzed by SDS-PAGE and subsequent Coomassie brilliant blue staining. (B) Activity assays using recombinant protein and 2  $\mu$ M of 5'-<sup>32</sup>P-labeled (indicated with an asterisk) RNA2 substrate were performed at 30°C for 2 or 20 min. The cleavage products were separated by electrophoresis in a 20% polyacrylamide/7 M urea gel and were visualized by phosphorimager analysis. A nucleotide ladder prepared by alkaline hydrolysis of RNA2 (AH marker), as well as 5'-<sup>32</sup>P-labeled synthetic short oligoribonucleotides (m2, m3, and m4), were included as size markers. Reactions without enzyme served as negative controls. The sequence of RNA2 is shown above the gel, and cleavage sites 3' of the cytidylates present in the first (■) and third (●) positions and 3' of the uridylylate in the sixth position (◆), and the corresponding cleavage products, are indicated.

*araBAD* promoter, whose generally low basal expression was further repressed by the addition of glucose to the bacterial growth medium prior to induction of nsp11 expression. The combination of maximally repressed T7 RNA polymerase expression and induction of nsp11 expression at low temperature allowed us to produce and purify wild-type nsp11 from EAV, as well as PRRSV (Fig. 2A). Nevertheless, the yield of wild-type nsp11 remained extremely low, and its toxicity was again illustrated by the fact that mutagenesis of putative catalytic residues dramatically improved expression levels under otherwise identical conditions (data not shown). We have used the same protocol for the expression and purification of SARS-CoV nsp15, whose endoribonuclease activity was previously characterized in some detail<sup>3,5,13,16</sup>. This allowed us to perform comparative functional studies of the arterivirus and coronavirus NendoU-containing nsps purified under identical conditions.

In order to assess their ability to process RNA *in vitro*, recombinant EAV and PRRSV nsp11 proteins were incubated with a 5'-<sup>32</sup>P-labeled synthetic oligoribonucleotide, RNA2

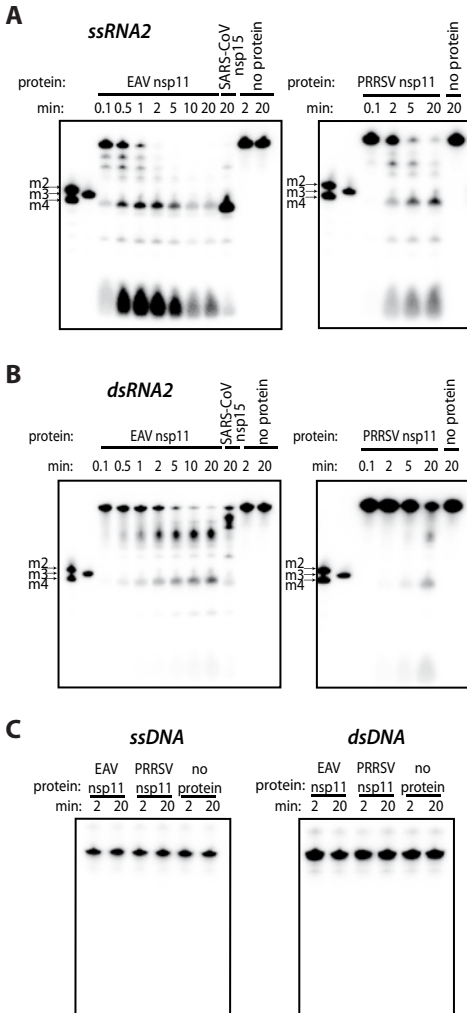
(5'-CGCAGUGAGCUCCUAAUCGCC-3'), that was shown to be efficiently cleaved by coronavirus NendoU proteins<sup>16</sup>. Both arterivirus proteins were able to cleave the substrate in the absence of divalent cations, and the resulting products were similar to those derived from cleavage of RNA2 by SARS-CoV nsp15 (Fig. 2B). The size of the major cleavage product of the SARS-CoV enzyme was consistent with processing 3' of the uridylylate present at the sixth position, although its migration in denaturing PAGE was altered due to the presence of an additional phosphate group at the 3' end that is lacking in the synthetic marker with the same nucleotide composition, RNA m3<sup>16,29</sup>. A product of identical size was generated by cleavage of RNA2 by both EAV and PRRSV nsp11 (Fig. 2B). The arterivirus enzymes also generated products consistent with processing at the two cytidylates present at the first and third positions in RNA2, which were detectable in the SARS-CoV NendoU-driven reaction, as well, albeit to a lesser extent. An RNA ladder of fragments produced by alkaline hydrolysis of RNA2, and therefore containing 3' phosphates, was used to verify the sizes of the cleavage products (Fig. 2B). These results indicated that arterivirus nsp11 possesses endonuclease activity and, in cleaving 3' of pyrimidines, exhibits an *in vitro* substrate specificity similar to that of SARS-CoV nsp15. Interestingly, the arterivirus enzymes seem to cleave at cytidylates more efficiently than the SARS-CoV NendoU, shown by the higher efficiency of processing at the cytidylate present at the first position in RNA2 (Fig. 2B), suggesting a less strict preference for uridylylate than previously reported for the coronavirus endoribonuclease<sup>4,16</sup>.

Prior studies had established that Mn<sup>2+</sup> ions enhanced the NendoU activity of coronavirus nsp15 *in vitro*<sup>5,16,38</sup> and were strictly required for RNA cleavage by XendoU<sup>11,19</sup>. The stimulating effect of Mn<sup>2+</sup> ions on SARS-CoV nsp15-mediated processing of RNA2 was confirmed in our assay (Fig. 3). In contrast, both EAV and PRRSV nsp11 proteins were able



**Figure 3.** Influence of Mn<sup>2+</sup> ions on arterivirus endoribonuclease activity. Recombinant proteins were incubated with 5  $\mu$ M of 5'-<sup>32</sup>P-labeled (\*) RNA2 at 30°C in the absence of Mn<sup>2+</sup> or in the presence of 1 mM or 5 mM of Mn<sup>2+</sup>, as indicated above the gel. The reactions were stopped after 2 or 20 min, and the cleavage products were separated by denaturing PAGE. For size markers and cleavage products, see the legend to Fig. 2.

to process RNA2 efficiently in reaction buffers lacking divalent cations (Fig. 2B), suggesting that metal ions are not required for catalysis. To address this potential discrepancy, we investigated the role of  $Mn^{2+}$  ions for arterivirus NendoU activity. Addition of  $Mn^{2+}$  ions at concentrations previously shown to stimulate coronavirus NendoU activity in fact inhibited the cleavage of RNA2 by both EAV and PRRSV nsp11, as was evident from the increase of remaining full-length substrate, as well as the accumulation of processing intermediates (Fig 3). Inhibition of arterivirus nsp11-mediated RNA cleavage in the presence of  $Mn^{2+}$  ions was consistently reproduced with a range of substrates (data not shown). These results led us to conclude that  $Mn^{2+}$  ions are not required for optimal NendoU activity of arterivirus nsp11 *in vitro*, and consequently, the subsequent functional characterization of the arterivirus enzymes was performed in the absence of  $Mn^{2+}$ .



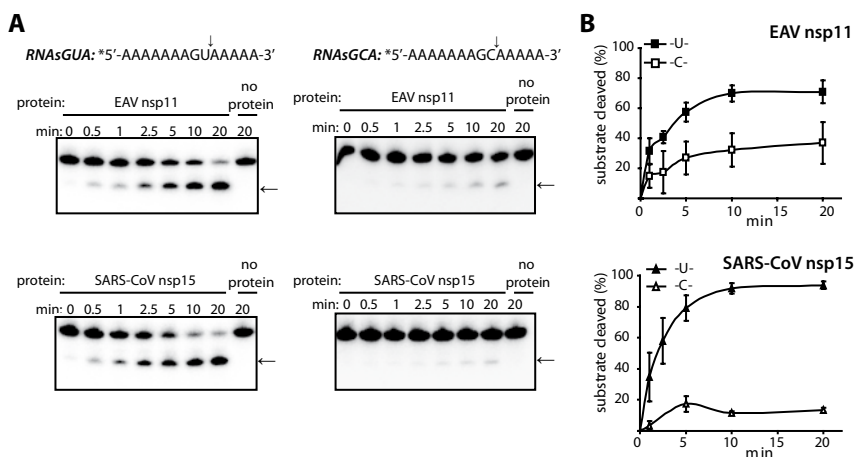
**Figure 4.** Substrate requirements of the arterivirus endoribonuclease. Activity assays containing recombinant protein and  $2 \mu M$  of  $5'$ - $^{32}P$ -labeled ssRNA2 (A), dsRNA2 (B), or ss/dsDNA (C) were incubated at  $30^\circ C$ , and aliquots were taken at the time points indicated above the lanes. For size markers, see the legend to Fig. 2.



## Arterivirus NendoU cleavage: substrate requirements and preference for uridylylate

To gain more insight into the substrate requirements of the arterivirus NendoU, we examined its ability to cleave ssRNA and dsRNA molecules *in vitro*. Both ssRNA2 and dsRNA2 were cleaved by the arterivirus NendoU. ssRNA molecules were processed more efficiently by both arterivirus nsp11 and SARS-CoV nsp15 (Fig. 4, A and B). Neither ssDNA nor dsDNA were hydrolyzed by the enzymes (Fig. 4C), confirming their specificity for RNA substrates.

Recent studies of SARS-CoV and MHV nsp15 have shown that they can both process RNA at cytidylates, although much less efficiently than at uridylylates<sup>4,18</sup>. The data we obtained using RNA2 as a substrate confirmed the strong preference of the SARS-CoV NendoU for uridylylate and indicated that the EAV endoribonuclease is able to process RNA efficiently at pyrimidines (Fig. 2B). The large number of uridylylates and cytidylates present in our initial substrate complicated the assessment of (potential) low efficiency cleavage events. Therefore, in order to establish whether EAV NendoU prefers uridylylate or cytidylate substrates, we used two additional oligoribonucleotides, both containing a single pyrimidine in an identical backbone: 5'-AAAAAAGUAAAAA-3' (RNAsGUA) and 5'-AAAAAAGCAAAAA-3' (RNAsGCA). Incubation of these substrates with EAV nsp11 or SARS-CoV nsp15 yielded a single cleavage product, consistent with the pyrimidine specificity of these enzymes (Fig. 5A). Quantification of the processing efficiency of these molecules under identical reaction conditions showed that both the EAV and the SARS-CoV NendoU proteins hydrolyzed the RNAsGUA substrate more efficiently than RNAsGCA (Fig. 5B). However, a larger

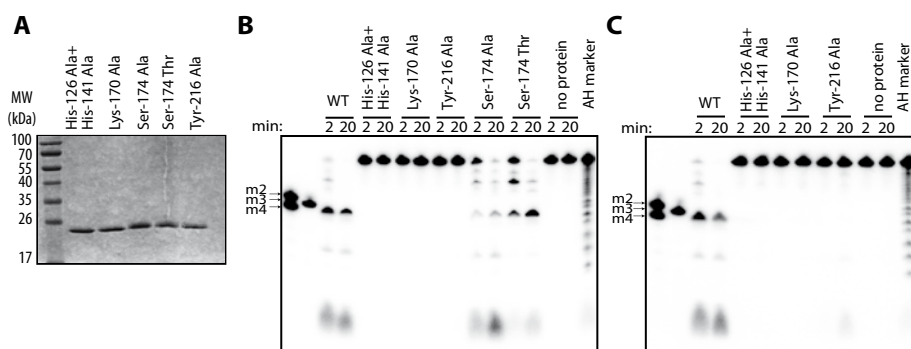


**Figure 5.** Arterivirus nsp11 cleaves RNA preferentially at uridylylates. (A) One micromolar 5'-<sup>32</sup>P-labeled synthetic oligoribonucleotide containing a single uridylylate (RNAsGUA) or a single cytidylate (RNAsGCA) was incubated with 0.3  $\mu$ M SARS-CoV nsp15 or 0.6  $\mu$ M EAV nsp11 at 30°C under standard reaction conditions. Aliquots were taken at different time points, and the reaction products were analyzed by denaturing PAGE and phosphorimager analysis. Cleavage sites and the corresponding products are indicated with arrows. (B) The amount of full-length substrate remaining was quantified, and the percentage of cleaved substrate was plotted as a function of the reaction time. The data points represent mean values from three independent experiments, and the error bars denote standard deviations.

fraction of the RNAsGCA substrate was cleaved by EAV nsp11 than by SARS-CoV nsp15 during the course of the reaction (Fig. 5B). These results suggest that EAV NendoU exhibits pyrimidine specificity with a clear preference for cleavage at uridylates, which nevertheless seems to be less pronounced than that of the SARS-CoV enzyme.

### Substitution of conserved amino acids in the EAV NendoU domain: effects on RNA cleavage and substrate specificity

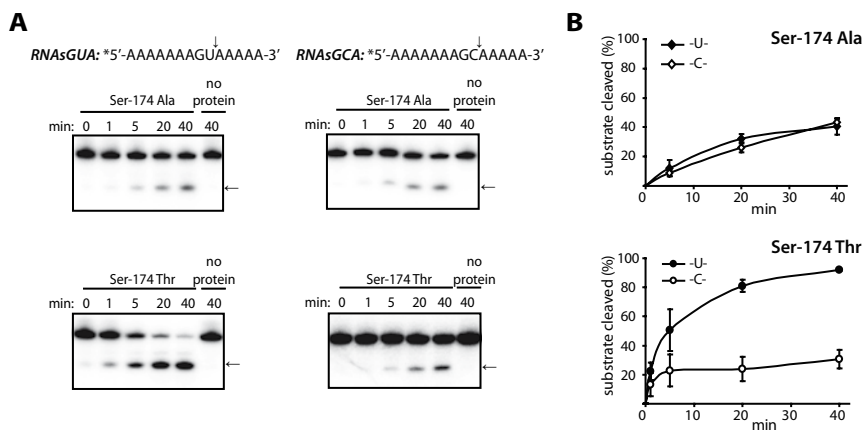
Functional studies of both XendoU and coronavirus NendoU proteins have confirmed the critical role of two His and a Lys residue in catalysis<sup>11,16,38</sup>. These three amino acids were initially proposed to form a catalytic center based on their conservation across the entire protein family<sup>33</sup>. Structural studies substantiated this hypothesis, showing that these residues cluster in a positively-charged groove of the SARS-CoV enzyme<sup>29</sup> and form a phosphate-binding site in XendoU<sup>28</sup>. Superimposing the spatial arrangements of the catalytic triads of SARS-CoV NendoU and RNase A revealed that two additional residues in SARS-CoV NendoU, Tyr-342 and Ser-293, interact with the substrate and possibly confer uridylate specificity to the enzyme<sup>29</sup>. To address the importance of the corresponding residues (indicated in the alignment in Fig. 1C) for the NendoU activity of EAV nsp11, they were replaced with Ala. The N-terminally His-tagged mutant proteins were expressed and purified under conditions identical to those used for purification of the wild-type nsp11 (Fig. 6A). A double replacement of His-126 and His-141, as well as single Lys-170- or Tyr-216 to Ala substitutions reduced cleavage of RNA2 to undetectable levels (Fig. 6B). Processing of this substrate was not observed even at enzyme concentrations 5-fold higher than those optimal for the wild-type protein (Fig. 6C), except for the Tyr-216 to Ala mutant, for which trace amounts of cleavage products were observed upon longer reaction times. Taken together, these data confirmed the critical role of the mutated residues in arterivirus NendoU-mediated RNA processing.



**Figure 6.** Catalytic activities of EAV nsp11 mutants. (A) Purified N-terminally His<sub>6</sub>-tagged EAV nsp11 proteins containing the indicated amino acid substitutions were electrophoresed on an SDS-polyacrylamide gel, which was subsequently stained with Coomassie brilliant blue. (B and C) Reaction mixtures containing 0.3 μM (B) or 1.5 μM (C) of mutant recombinant proteins and 2 μM of 5'-<sup>32</sup>P-labeled RNA2 were incubated at 30°C. The reactions were stopped at the indicated time points, and the cleavage products were separated in a 20% polyacrylamide/7 M urea gel. For size markers, see the legend to Fig. 2. WT, wild type.

The Ser-174 to Ala substitution also impaired EAV NendoU activity, although not as dramatically, since cleavage of RNA2 was clearly detectable at a protein concentration similar to that used for wild-type EAV NendoU (Fig. 6B). Comparing the processing efficiency of RNAsGUA versus RNAsGCA revealed that the Ser-174 to Ala substitution had rendered the EAV NendoU unable to discriminate between uridylates and cytidylates, since the mutant protein was able to cleave at both nucleotides with a similar efficiency (Fig. 7A and B). We also replaced Ser-174 with Thr, a residue that is present at this position of the NendoU alignment in all other nidovirus homologs, with the exception of SARS-CoV and closely related bat coronaviruses (see Fig. S1 in the supplemental material). This substitution resulted in somewhat impaired *in vitro* RNA processing in comparison with the wild-type EAV NendoU, as well (Fig. 6B), but it had no effect on the uridylate preference of the enzyme (Fig. 7A and B). These results strengthen the view that Ser-174 is important for substrate recognition and is involved in determining the substrate specificity of the EAV NendoU.

Our previous analysis of the impact of NendoU mutagenesis on the EAV life cycle revealed that substitution of His-126, His-141, Lys-170 or Ser-174 with Ala caused major defects in viral RNA synthesis and production of infectious progeny<sup>24</sup>. In addition, the critical role of two nidovirus-wide conserved aspartic acid residues, Asp-177 and Asp-201 (Fig. 1C) in viral RNA synthesis was established. Non-conservative (to Ala) or conservative (to Glu or Asn) substitutions of those residues rendered viral replication undetectable. Examining the effect of these mutations on the *in vitro* endonuclease activity of EAV NendoU could therefore provide valuable insight into the basis of this nonviable phenotype. Our attempts to purify recombinant EAV nsp11 containing an Asp-177 to Ala or Asp-201 to Ala substitution, however, were unsuccessful, since both mutations

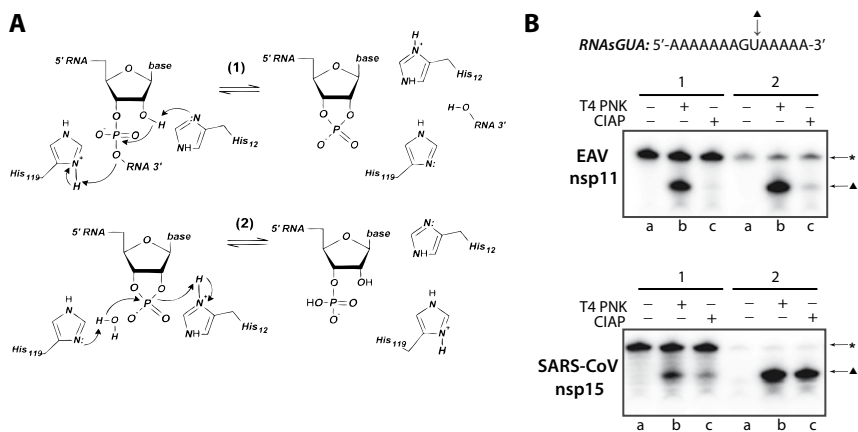


**Figure 7.** Mutagenesis of Ser-174 influences EAV nsp11 substrate specificity. (A) One micromolar of RNAsGUA or RNAsGCA was incubated with 0.6  $\mu$ M EAV nsp11 S3011A or 0.6  $\mu$ M EAV nsp11 S3011T at 30°C under standard reaction conditions. Aliquots were taken at different time points, and the reaction products were analyzed by denaturing PAGE and phosphorimager analysis. (B) The amount of full-length substrate remaining was quantified, and the percentage of cleaved substrate was plotted as a function of the reaction time. The data points represent mean values from three independent experiments, and the error bars denote standard deviations.

rendered the protein insoluble. Similar results were reported following mutagenesis of the Asp-177 counterpart in SARS-CoV NendoU<sup>13</sup> and the Asp-201 equivalent in the MHV enzyme<sup>18</sup>. The crystal structure of SARS-CoV nsp15 revealed that these residues are part of an extensive hydrogen bond network in close proximity to the active site and that replacing them could have an indirect effect on the activity of the protein by disturbing its secondary structure<sup>29</sup>, a notion supported by the insolubility of the recombinant mutant proteins.

### **EAV NendoU RNA cleavage: reaction mechanism**

A number of parallels have been drawn between the active sites of RNase A, XendoU<sup>28</sup>, coronavirus NendoU<sup>29</sup> and, in this report, arterivirus NendoU. RNase A is a remarkably well characterized enzyme that has been proposed to cleave the P-O<sup>5'</sup> bond in RNA via general acid-base catalysis, involving the side chains of the active site His-12 and His-119 residues (reference<sup>26</sup> and references therein). In a first reaction, transphosphorylation of the RNA results in the accumulation of 2',3'-cyclic phosphodiester, which can be subsequently hydrolyzed to 3'-phosphomonoesters in a separate process<sup>34</sup> (Fig. 8A). 2',3'-Cyclic phosphates have been detected at the 3' ends of cleavage products generated by both XendoU<sup>19</sup> and coronavirus NendoU<sup>16</sup>, suggesting similarities to RNase A in terms of cleavage chemistry. To investigate the chemical nature of the 3'-termini of cleavage products generated by the EAV NendoU, the products derived from cleavage of RNAsGUA after short (2-min) or prolonged (40-min) reaction times were treated with T4 PNK or CIAP. Following phenol extraction and ethanol precipitation, the cleavage products were ligated to [5'-<sup>32</sup>P]pCp in a T4 ligase-mediated reaction. As expected, pCp could be readily ligated to the hydroxyl group present at the 3' end of the unprocessed substrate (Fig. 8B), and this reaction was used as an internal control for determining the efficiency of pCp ligation. In contrast to the uncleaved substrate RNA, efficient transfer of the [<sup>32</sup>P] label to the 3' terminus of the 5' cleavage product obtained after incubation with EAV nsp11 required phosphatase treatment prior to the pCp ligation reaction, suggesting the presence of a 3' phosphate group. After a short EAV NendoU reaction, the 5' cleavage products could be efficiently ligated only after dephosphorylation with T4 PNK (which is capable of removing both 2',3' as well as 3' phosphates) but not after CIAP treatment (that can remove only 3' phosphates), which suggests that the 5' cleavage products contained predominantly 2',3'-cyclic phosphates. Upon longer incubation of substrate and enzyme, however, CIAP treatment was sufficient for efficient ligation of a portion of the 5' cleavage products to [5'-<sup>32</sup>P]pCp, indicating that nsp11-mediated hydrolysis of part of the cyclic phosphodiester had occurred (Fig. 8B). In the SARS-CoV nsp15-mediated reaction, the fraction of 5' cleavage products whose cyclic phosphate ends had been hydrolyzed was even larger, as illustrated by the higher efficiency of [5'-<sup>32</sup>P]pCp ligation after dephosphorylation with CIAP (Fig. 8B). These results suggest that the EAV and SARS-CoV NendoU's are able to catalyze both transphosphorylation of RNA to form a 2',3'-cyclic phosphodiester, as well



**Figure 8.** Analysis of the 3' termini of EAV and SARS-CoV NendoU products. (A) Proposed mechanism for the transphosphorylation (1) and hydrolysis (2) reactions catalyzed by RNase A. (B) Cleavage products obtained after incubation of 1  $\mu\text{M}$  unlabeled RNAsGUA with 0.6  $\mu\text{M}$  EAV nsp11 or SARS-CoV nsp15 for 2 min (1) or 40 min (2) at 30°C under standard reaction conditions were subsequently ligated to 5'- $^{32}\text{P}$ pCp without further treatment (lanes a) or after dephosphorylation with T4 PNK (lanes b) or CIAP (lanes c). The reaction products were analyzed by denaturing PAGE and phosphorimager analysis. The full-length substrate (\*) and the 5' cleavage product (▲) are indicated.

as its subsequent hydrolysis to a 3'-phosphomonoester, which would be consistent with an RNase A-like reaction mechanism.

## DISCUSSION

### The NendoU activity is common to arteri- and coronaviruses, and likely all nidoviruses.

This report describes the *in vitro* endoribonuclease activity of nsp11 from two quite distantly related arteriviruses, EAV and PRRSV, and expands the short list of characterized XendoU/NendoU superfamily members with representatives of a new, deeply separated phylogenetic lineage. The sizes and domain compositions of the nidovirus replicase subunits harboring NendoU domains vary considerably among nidoviruses, which might be linked to differential catalytic properties and functions of the enzymes in the life cycle of different nidoviruses. The biochemical characterization of arterivirus nsp11 documents its ability to hydrolyze RNA 3' of pyrimidines, with a preference for uridylates, and revealed that this endoribonuclease activity is not dependent on the presence of divalent cations. Mutagenesis of predicted active site residues abolished the protein's ability to cleave RNA and provided insights into the determinants of the enzyme's uridylate preference. Our comparative analysis of arterivirus nsp11 and coronavirus nsp15 purified under identical conditions established important parallels between the ribonucleolytic activities of these enzymes, such as the identity of key residues in their catalytic site

and their proposed reaction mechanism, which are likely conserved across the entire nidovirus order. Our data also revealed differences in their biochemical properties, such as the effect of  $Mn^{2+}$  on their *in vitro* activity, as well as their variable extent of preference for cleavage at uridylates, both probably owing to considerable divergence of arterivirus nsp11 and coronavirus nsp15, which carry distantly related NendoU's fused to apparently unrelated N-terminal domains (Fig. 1C).

The endoribonuclease activity of EAV and PRRSV nsp11 seems to be independent of the presence of metal ions as cofactors, since both proteins cleaved RNA very efficiently in reaction buffers lacking divalent cations and containing EDTA (Fig 2B, 3). This finding stands in contrast to the previously documented stimulatory effect of  $Mn^{2+}$  on NendoU activity of coronavirus nsp15 *in vitro*<sup>16,38</sup>, which was corroborated by the comparative enzyme activity studies included in this paper (Fig. 3). Our data show that arterivirus endoribonucleolytic activities are inhibited by  $Mn^{2+}$  ions at concentrations that have previously been shown to significantly enhance the activities of a number of coronavirus NendoU enzymes<sup>3,16,38</sup>. Interestingly, similar  $Mn^{2+}$  concentrations have been reported to inhibit RNase A-mediated RNA hydrolysis as well, probably by binding to and blocking of the catalytic histidine residues<sup>1</sup>. The same study showed that low concentrations of  $Mn^{2+}$  may enhance RNase A activity, possibly through interactions between the metal ions and the phosphates present in the RNA substrate. Due to the fast rate of RNA cleavage by the recombinant enzymes and the gel electrophoresis-based activity assays used in this study, we have been unable to determine if arterivirus *in vitro* endoribonuclease activities can be further stimulated by low concentrations of  $Mn^{2+}$  or other divalent cations. It should be noted that the precise role of  $Mn^{2+}$  in RNA hydrolysis by the coronavirus NendoU and its cellular homologs, XendoU and human PP11, remains enigmatic. Although  $Mn^{2+}$  is essential for *in vitro* RNA processing by XendoU and human PP11, the interaction of the enzymes with a substrate does not require metal ions<sup>11,19,20</sup>. The role of  $Mn^{2+}$  for coronavirus NendoU cleavage is even less clear. Previous studies have implicated  $Mn^{2+}$  ions in triggering conformational changes<sup>3</sup>, as well as affecting RNA binding properties of these proteins<sup>5</sup>. Metal binding sites have not been identified in the currently available crystal structures of coronavirus nsp15<sup>4,17,29,38</sup> or XendoU<sup>28</sup>. Furthermore, millimolar concentrations of  $Mn^{2+}$  ions are required for enhancing RNA cleavage by coronavirus nsp15 *in vitro*, while the intracellular free  $Mn^{2+}$  concentration is estimated to be in the micromolar range (references<sup>5,27</sup> and references therein), raising questions about the biological significance of this observation. Comparative analysis of both the structure and biochemistry of the arteri- and coronavirus enzymes could therefore be used to gain more insight into the specific role(s) of divalent cations in coronavirus NendoU-mediated cleavage.

### **The EAV NendoU active site and mechanism of RNA cleavage**

Replacement of the predicted catalytic residues (His-126, His-141 and Lys-170) of EAV NendoU with alanine established their importance for *in vitro* RNA processing (Fig. 6, B and C). This observation is consistent with the alignment in Fig. 1C and results of similar

experiments with related enzymes of other origins, as well as RNase A. Likewise, the substitution of Tyr -216, whose counterpart in the coronavirus NendoU was implicated in the proper positioning and orientation of the substrate in the coronavirus NendoU active site<sup>29</sup>, also severely impaired EAV NendoU-mediated cleavage *in vitro* (Fig. 6B and C). Interestingly, a replacement of Ser-174 with Ala only moderately affected cleavage by the EAV NendoU (Fig. 6B), but rendered the enzyme unable to discriminate between uridylylates and cytidylylates (Fig. 7). The conservative replacement of Ser-174 with Thr largely restored efficient RNA processing, including the preference of EAV NendoU for cleavage at uridylylates (Fig. 7). Ser-174 of EAV nsp11 is the counterpart of residue Ser-293 in SARS-CoV nsp15<sup>29</sup>, which has been proposed to be functionally equivalent to Thr-45 of RNase A, a major determinant of this enzyme's pyrimidine specificity (reference<sup>26</sup> and reference therein). In a recent study<sup>4</sup>, the effects of SARS-CoV nsp15 Ser-293 substitutions by Ala or Thr were reported to be similar to the ones we describe here for EAV nsp11 Ser-174 replacements. Two additional residues of SARS-CoV nsp15, Pro-343 and Leu-345, which are highly conserved in the C-terminal domain of coronavirus NendoU domains (Fig. 1C), were also implicated in determining the substrate specificity of the endoribonuclease. Pro-343 was proposed to sterically hinder the access of cytosine to the active site, while the side chain of Leu-345 could prevent the access of purines. The extreme C-terminal residues of the EAV and SARS-CoV NendoU proteins differ substantially, with the counterpart residue of Pro-343 being Val-217 in EAV nsp11, while the bulky Leu-345 presumably has no counterpart (see alignment in Fig. 1C). Structural information on the arterivirus endoribonuclease would shed more light on whether the protein's C-terminal tail is also involved in substrate binding and whether the limited conservation of this region could account for the more relaxed uridylylate specificity we observed for the EAV NendoU in comparison to the SARS-CoV enzyme (Fig. 5).

The similarities between the active sites of viral NendoU proteins and RNase A, which were previously suggested by others<sup>29</sup> and are also outlined in this report, prompted us to examine the mechanism of RNA cleavage by the EAV endoribonuclease. The reaction products of coronavirus NendoU were shown to possess 2',3'-cyclic phosphate ends<sup>16</sup>. Analysis of the 3'-ends of cleavage products generated by the EAV NendoU indicated that also this reaction proceeds through the formation 2',3'-cyclic phosphodiester, a fraction of which appear to be hydrolyzed to 3' phosphomonoesters upon longer reaction times (Fig. 8B). We obtained similar results upon analysis of SARS-CoV NendoU cleavage products. These data are consistent with the utilization of an RNase A-like mechanism of RNA cleavage by both the arteri- and coronavirus endoribonucleases. The ability of the SARS-CoV enzyme to hydrolyze 2',3'-cyclic phosphodiester had not been detected previously, most likely due to the relatively low activity of the MBP-nsp15 fusion protein used in these earlier studies.

## Role of NendoU in the life cycle of arteri- and coronaviruses

Despite recent advances in our understanding of the biochemical properties of nidovirus RNA-processing enzymes, the characterization of their precise roles and mechanisms of action in the viral life cycle is still at an initial stage. Reverse genetics-based mutagenesis studies of the coronavirus nsps that harbor the exoribonuclease (ExoN) and NendoU domains confirmed their importance for coronavirus RNA synthesis<sup>7,8,16,18,21</sup>, and lent support to the involvement of ExoN in the fitness control of coronavirus RNA replication<sup>8</sup>. A site-directed mutagenesis study of EAV nsp11 in the context of the complete EAV life cycle suggested that this replicase subunit is a multifunctional viral protein<sup>24</sup>. Substitution of Asp-177 or Asp-201, which in this report were found to render bacterially expressed EAV nsp11 insoluble, resulted in a non-viable phenotype. A similar result was obtained by deleting the nsp11-coding sequence from the EAV genome. Most notably, however, replacement of any of the three putative catalytic residues (His-126, His-141, or Lys-170), which were shown in this study to reduce the *in vitro* EAV NendoU activity to background levels, yielded severely crippled, but viable mutant viruses. These substitutions induced a specific defect in viral sg mRNA synthesis, which was accompanied by a dramatic reduction of the production of infectious progeny. We cannot formally exclude that a basal level of endoribonuclease activity, undetectable in our *in vitro* assay, may be retained by these catalytic site mutants *in vivo*, which might be sufficient for viral replication and production of virus progeny, albeit at a low level. Examination of the role of the MHV NendoU in the viral life cycle revealed that mutagenesis of the catalytic His residues resulted in severely impaired, but detectable RNA processing activity *in vitro*, and only modest defects in viral RNA synthesis and production of infectious progeny<sup>18</sup>. Furthermore, no specific defects in sg mRNA synthesis were detected. This stands in contrast to the observations made upon EAV NendoU mutagenesis by reverse genetics<sup>24</sup> and results obtained by replacement of an NendoU active-site His in a SARS-CoV replicon. In the latter case, a severe decrease of sg mRNA levels was observed<sup>2</sup>. Overall, these data suggest (partially) different roles of NendoU in the life cycles of arteri- and coronaviruses. Alternatively, a potentially higher basal activity retained by the MHV NendoU catalytic site mutants, combined with differences in overall replication rates between MHV and EAV, might minimize the consequences of these substitutions for the MHV life cycle. Moreover, the significance of the uridylyate preference observed for arteri- and coronavirus endoribonucleases *in vitro* for viral RNA synthesis is difficult to assess; the *in vivo* substrates of these enzymes are still unknown, and identifying them will be a key step in linking the available *in vitro* and *in vivo* data mechanistically. In this respect, it remains unclear whether the reduction of EAV sg mRNA synthesis and the decrease in infectious progeny production seen upon substituting Ser-174 of nsp11 with Ala<sup>24</sup> can be attributed to the loss of the mutant enzyme's ability to discriminate between uridylyates and cytidylyates (Fig. 7) or to its overall lower activity (Fig. 6B). In addition, the importance of viral determinants of pathogenesis can be difficult to assess in tissue culture, as illustrated by the coronavirus ADP-ribose 1''-phosphatase (ADRP) and the putative cyclic



phosphodiesterase (CPD) domains, which are dispensable for replication in cultured cells, but critically involved in viral pathogenicity *in vivo*<sup>10,25,30</sup>

Irrespective of the exact function of arteri- and coronavirus NendoU proteins in the viral life cycle, the broad specificity that these endoribonucleases exhibit *in vitro* suggests that their activity in infected cells must be strictly regulated in order to protect viral (and possibly cellular) RNA species from rapid degradation. Indeed, expression of EAV nsp11 outside the context of the infected cell is extremely toxic to prokaryotic and eukaryotic cells (data not shown). For example, protein-protein interactions or modulation by cofactors might act as allosteric switches to regulate the ribonucleolytic activity of arterivirus nsp11 and coronavirus nsp15, and prevent nonspecific cleavage of viral and/or cellular RNAs. The hexamerization of coronavirus nsp15, mediated by its N-terminal domain, and the presence of Mn<sup>2+</sup> have been implicated in ensuring optimal RNA processing activity<sup>5,13,17</sup>. By contrast, EAV NendoU activity is inhibited in the presence of Mn<sup>2+</sup> (Fig. 3) and its quaternary structure may differ as well. Although the yield of bacterially expressed EAV nsp11 was unfortunately too low to attempt to analyze its oligomeric state, preliminary gel filtration analyses of catalytic site mutants suggest they are present as monomers in solution (data not shown). This observation may turn out to be linked to the properties of the N-terminal domain of arterivirus nsp11, which is relatively small and apparently unrelated to its coronavirus nsp15 counterpart (Fig. 1C) (A.E. Gorbalenya, unpublished). In conclusion, the spatial and temporal regulation of EAV NendoU activity is likely ensured by coordinated replicase polyprotein cleavage, interactions with other replicase subunits or cellular proteins, and/or its compartmentalization in a specifically organized membrane-bound replication/transcription complex. The continued biochemical and structural characterization of nidovirus NendoU's, in combination with the functional dissection of NendoU-containing ribonucleoprotein complexes in infected cells, should provide a basis for the identification of the *in vivo* target(s) of this replicative endoribonuclease that is unique in the RNA virus world.

## ACKNOWLEDGEMENTS

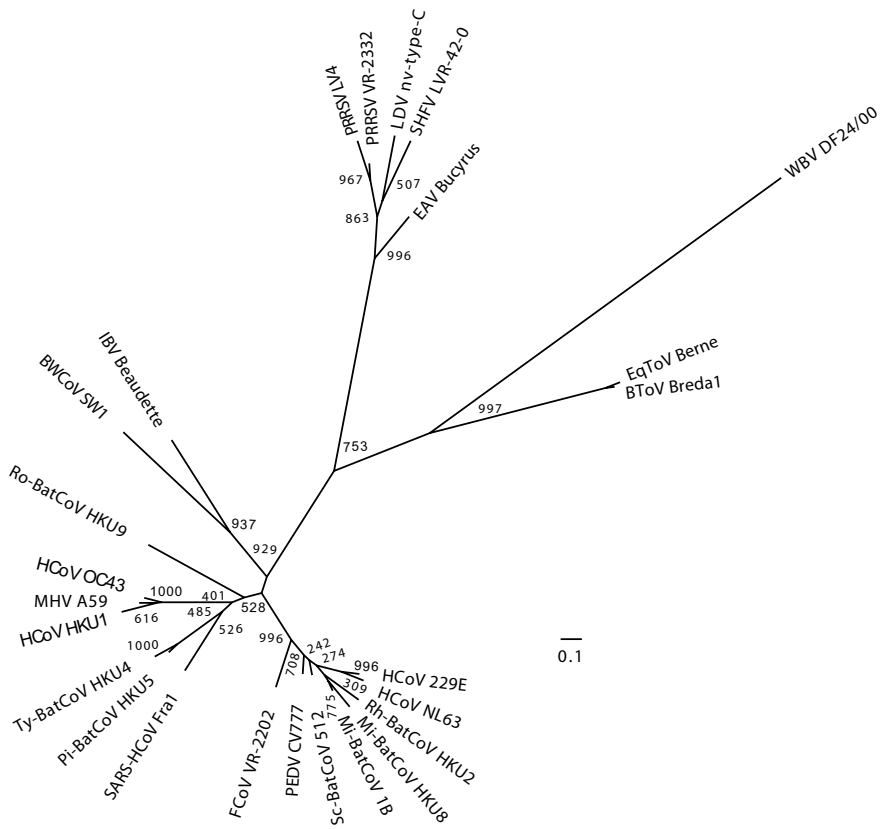
We are grateful to our LUMC colleagues Linda Boomaars-van der Zanden and Inge Schaap and to Bruno Canard, Bruno Coutard and their co-workers of the VIZIER expression platform (CNRS and University of Aix-Marseille, France) for generating the PRRSV nsp11 expression construct. A.E.G. thanks Michael Rozanov for sharing his observations. We thank Sjoerd van den Worm and Martijn van Hemert for helpful discussions, and Igor Sidorov and Alexander Kravchenko for help with ViraLis and its databases.

This work was supported by the Council for Chemical Sciences of the Netherlands Organization for Scientific Research (NWO-CW; grant 700.52.306) and the VIZIER integrated project (LSHG-CT-2004-511960) of the European Union's 6th Framework Programme.

# SUPPLEMENTAL MATERIAL

BToV Bredal	[NC_007447]	[1-142]	GNGESLRELPELTFSTGRLYN-----LEHDPKSNFNVOQLA-ETIPKRVHVFAGDFDVTGTDIGGVHVVVALNGYK-----
EqToV Berne	[X52374]	[1-142]	GGKESLFLDPTLFTSTGRLYN-----LDHDAQAQNFNVKQLA-ETPNMNVHVSQDFTEVGTGDIQGVHVVVALNGYK-----
WBV DF24/00	[NC_008516]	[1-135]	----FVHGQSTFDLGLYST-----NRLTKQFYTYPMDEG-VKSGKLVHVSQDTSNVSMTIGGHLVTFAPQNTYI
IBV Beaudette	[NC_001451]	[1-171]	VNQA-EVFLPNTVNTQGRVYFFPRSDIERDFLANSEESVEVRYG-KDGLGLILY-GEVSK--FDLGGHLLIQLRARIKANKL
BWCoV SW1	[EU111742]	[1-182]	-DGR-HVPTQSOLFTEQGRSFDLFLPSVNERDFLNLEAPQIAKYDCKGLGLVLY-GDFSK-KVIGGALITGLARIQEKDF
HCoV-NL63	[NC_005831.2]	[1-179]	-NGQ-FQDHYDGFYTGQRNLSDTFRSDMEYDFLNMMDGVFINKYGLEDFENFVIV-GDVK--TTLGGHLLISQERLSKGMVGLK
HCoV-229E	[NC_002645]	[1-183]	-DGR-PVDHYDGFYTGQRNLSDTFRSDMEYDFLNMMDGVFINKYGLEDFENFVIV-GDVK--TTLGGHLLISQERLSKGMVGLK
Mi-BatCoV 1B	[NC_010436]	[1-174]	-DGV-FVDCQDGIPTQGRNVISFFPRSDMEYDFLNMMDGVFINKYGLEDFENFVIV-GDVK--TTLGGHLLISQERLSKGMVGLK
Mi-BatCoV HKU8	[NC_010438]	[1-174]	-DGV-FVDCQDGIPTQGRNVISFFPRSDMEYDFLNMMDGVFINKYGLEDFENFVIV-GDVK--TTLGGHLLISQERLSKGMVGLK
Rh-BatCoV HKU2	[NC_009988]	[1-178]	-DGS-FVDPVQDYSIQGRNASTFLPRSQMEKDFLELDGLPIKSYGLEDFENFVIV-GDVK--TTLGGHLLISQERLSKGMVGLK
Sc-BatCoV 512	[DQ648858]	[1-174]	-DGA-FVEYDQDYSIQGRNVISDFQPRSDMEYDFLNMMDGVFINKYGLEDFENFVIV-GDVK--TTLGGHLLISQERLSKGMVGLK
PEDV CV777	[NC_003436]	[1-174]	-NGK-FEDYDQDYSIQGRNVISDFQPRSDMEYDFLNMMDGVFINKYGLEDFENFVIV-GDVK--TTLGGHLLISQERLSKGMVGLK
FCoV VR-2202	[NC_007025]	[1-174]	-NGE-FVQDQDYSIQGRNVISDFQPRSDMEYDFLNMMDGVFINKYGLEDFENFVIV-GDVK--TTLGGHLLISQERLSKGMVGLK
MHV-A59	[NC_001846]	[1-209]	LSGN-EALARGTIFTPQSRLLSFTFPRSDMEYDFLNMMDGVFINKYGLEDFENFVIV-GSFNQ--KIIGGLHLLIGLARQQRKSNLV
HCoV-HKU1	[NC_006677.2]	[1-209]	VIGN-DALTRFTTIPQSRVLSFSEPRSDLRDFIDMDDNLPKIAKYLQDVAFFVIV-GSFNH--KVIIGGLHLLIGLFRKRRKSNLL
HCoV-OC43	[NC_005147]	[1-210]	VGN-DALISITFTQSRVLSFTCTRDMKDFIALDQVDFIQKYGLEDVAFVIV-GFNQ--KIIIGGLHLLIGLRQRQTSNLV
Ty-BatCoV HKU4	[EP665505]	[1-177]	VNNE-FVFSDFTYGCRVTDVFTLPMEDFLVLDSDVFIKRYGLEDFENFVIV-GDVSF--TTLGGHLLIGLKRKRGHLL
Pi-BatCoV HKU5	[EP665509]	[1-186]	VNNE-FVQFTDVTYGLRQVSDTTPVSEMEKDFIALSDVFIKRYGLEDFENFVIV-GDVSF--TTLGGHLLIGLKRKRGHLL
Ro-BatCoV HKU9	[EP665513]	[1-175]	-DGA-FVQFTDGTYSQGRVTDVDFQPTQLEIDFLDEQSGFLDKYKLVHLDLGLLEIVIV-GDVF--GTIGGLHLLIGLARVRRKTAHLV
SARS-HCoV Fra1	[AY291315]	[1-181]	VGDGI-TQQLPETYFTQSRVLDLDFKPRSQMEKDFLELDGLPIKSYGLEDFENFVIV-GDVSF--GOLGGHLLIGLARVRRKTAHLV
PRRSV VR-2332	[AV150564]	[1-92]	--GE-AQVLPETVFTQGRVTDVDFQPTQLEIDFLDEQSGFLDKYKLVHLDLGLLEIVIV-GDVF--GTIGGLHLLIGLARVRRKTAHLV
PRRSV LV4	[AY588319]	[1-93]	--GE-AQVLPETVFTQGRVTDVDFQPTQLEIDFLDEQSGFLDKYKLVHLDLGLLEIVIV-GDVF--GTIGGLHLLIGLARVRRKTAHLV
LDV nv-type-C	[L13298]	[1-92]	--GK-PVPLPDSLSMSTRGIALNV-----REYLDSEKMEPSS-----RHFAPL-GEVSK--SNVGGCVHVSRY-----
SHFV LVR-42-0	[NC_003092]	[1-93]	--GR-AVPMEDSVYSTRGFEMDI-----RDYLDSEKDFPAA-----KHFAPL-GEVSK--SNVGGCVHVSRY-----
EAV Bucyrus	[NC_002532.2]	[1-89]	--GR-ARALPDSLFSGRGFETNS-----RAFLDSEKFPAA-----AHFLACL-GDITK--STVGGCVHVSRY-----
BToV Bredal	[NC_007447]		-----GSIIPNVKPIATGLI--NVGRNR- <b>LDL</b> -----VNCANLYEKVQGVQ--IAGVKVSKVIFNVDFQFVQVFMFKAGGEDIGTFPPQKE
EqToV Berne	[X52374]		-----GSIIPNVKPIATGLI--NVGRNR- <b>LDL</b> -----VNCANLYEKVQGVQ--LEGVVKVIFNVDFQFVQVFMFKAGGEDIGTFPPQKD
WBV DF24/00	[NC_008516]		-----YQLVSGFANPIMRINV--ALEGRNVEI <b>ALD</b> TTLLQDYKIAITN--KVT--VSKTFTFDGGQYRMNFANPDGTFQTSVFAQ
IBV Beaudette	[NC_001451]		AKSVTNSSDVMQNVFVL-SDNGSVQVQVQVQV <b>QL</b> LLDDEFELELLNLIKKEYGNKSKVTVVDSHNSINFMVFE-DGSKIKTCVQLO
BWCoV SW1	[EU111742]		VHEISPSDFVKSFFVY-THAGAGVQV <b>QL</b> VLLDDFTSILKEIKVSKWQVSEVLIYIYQTVDFELMWCN-DANITFEPQL
HCoV-NL63	[NC_005831.2]		ADDFVTAASDTRLCCTCYTYLNLSSVQV <b>QL</b> LLDDEFEVILKSL--DLGV-ISKVHEVIVDKVFRWMLWCK-DNHLSFTFPQL
HCoV-229E	[NC_002645]		AEEFVAASDTRLCCTCYTYLNDPSSVQV <b>QL</b> LLDDEFEVILKSL--DLTV-ISKVHEVIVDKVFRWMLWCK-DNAVATFPQL
Mi-BatCoV 1B	[NC_010436]		VEDFVSSDSTLKSCTVYNDPSSVQV <b>QL</b> LLDDEFEVILKSL--DLSV-ISKVHEVIVDKVFRWMLWCK-DYKQVTFPQL
Mi-BatCoV HKU8	[NC_010438]		VEEFVHTGDSLRCAVYTYNDPSSVQV <b>QL</b> LLDDEFEVILKSL--DLSV-ISKVHEVIVDKVFRWMLWCK-DYKQVTFPQL
Rh-BatCoV HKU2	[NC_009988]		VEDFVSSDSTLKSCTVYNDPSSVQV <b>QL</b> LLDDEFEVILKSL--DLSV-TSKVHVIVDKVFRWMLWCK-DKVAATFPQL
Sc-BatCoV 512	[DQ648858]		VEDFVSSDSTLKSCTVYNDPSSVQV <b>QL</b> LLDDEFEVILKSL--DLSV-ISKVHEVIVDKVFRWMLWCK-DYKQVTFPQL
PEDV CV777	[NC_003436]		IDEFVSSDSTLKSCTVYADNPSSVQV <b>QL</b> LLDDEFEVILKSL--DLSV-ISKVHEVIVDKMWRMLWCK-DHKLQTFPQL
FCoV VR-2202	[NC_007025]		VEFMFNSDSTLKSCTVYADNPSSVQV <b>QL</b> LLDDEFEVILKSL--DLNV-ISKVVDVIVDKANRWMLWCK-NSQIKTFPQL
MHV-A59	[NC_001846]		IQEFVT-YDSIHSYFTDENSGSSVQV <b>QL</b> LLDDEFEVILKSL--NLCK-ISKVNVNVDFKDFQFMLWCN-EKKNITFEPQL
HCoV-HKU1	[NC_006677.2]		IQEFVQ-YDSIHSYFTDENSGSSVQV <b>QL</b> LLDDEFEVILKSL--NLCK-ISKVNVNVDFKDFQFMLWCN-EKKNITFEPQL
HCoV-OC43	[NC_005147]		VEEFVY-YDSIHSYFTDENSGSSVQV <b>QL</b> LLDDEFEVILKSL--NLCK-ISKVNVNVDFKDFQFMLWCN-EKKNITFEPQL
Ty-BatCoV HKU4	[EP665505]		MEEMLK-DRAVTNHYFTDNTASVAVQV <b>QL</b> LLDDEFEVILKSL--DLTV-ISKVNVNVDFKDFQFMLWCN-EKKNITFEPQL
Pi-BatCoV HKU5	[EP665509]		MEEMLK-DRAVTNHYFTDNTASVAVQV <b>QL</b> LLDDEFEVILKSL--DLTV-ISKVNVNVDFKDFQFMLWCN-EKKNITFEPQL
Ro-BatCoV HKU9	[EP665513]		MEVFLQ-TD-VTSVAVDIDPQASVAVQV <b>QL</b> LLDDEFEVILKSL--DRSV-ISKVQCCVDFKDFQFMLWCN-EKKNITFEPQL
SARS-HCoV Fra1	[AY291315]		LEDFIP-NDSTVKNVFTDAQRGSSEVQV <b>QL</b> LLDDEFEVILKSL--DLSV-ISKVNVNVDFKDFQFMLWCN-EKKNITFEPQL
PRRSV VR-2332	[AV150564]		LRPVLV--KESVAVVGVG-SPGKAPAL <b>CL</b> LDVYLPDLRLAYLHPE---TQ--SKCKWMLDKFVRLMWKRD---KTAEPQLE
PRRSV LV4	[AY588319]		LRPVLV--KDSVAVVGVG-SPGKAPAL <b>CL</b> LDVYLPDLRLAYLHPE---TA--SKCKWMLDKFVRLMWKRD---KTAEPQLE
LDV nv-type-C	[L13298]		LRPVLV--PGSVKRVGVG-CPGKAPAL <b>CL</b> LDVYLPDLRLAYLHPE---TK--SMDYKLLDFKQKRLMWKRD---ATAEPH-E
SHFV LVR-42-0	[NC_003092]		LRPVLV--ADSVKRVGVG-SPGKAPAL <b>CL</b> LDVYLPDLRLAYLHPE---TQ--SKVYKMLDKFVRLMWKRD---KTAEPH-E
EAV Bucyrus	[NC_002532.2]		LRPVLV--ADAVALLVGS-LAGKAPAL <b>CL</b> LDVYLPDLRLAYLHPE---TL--SRVYKIMDKFVRLMWKRD---ATAEPH-E

**Figure S1.** Amino acid alignment of the NendoU domain of corona-, toro-, bafni- and arteriviruses. Virus acronyms are indicated in the first column. NCBI accession numbers of viral genomes and distances of the NendoU domain from the N-terminus of the protein (in amino acids) are listed in the second and third column, respectively. Conserved residues discussed in the text have been highlighted. Abbreviations are as follows: BToV Bredal1, bovine torovirus Breda-1; EqToV Berne, equine torovirus Berne; WBV DF24/00, white bream virus; IBV Beaudette, infectious bronchitis virus strain Beaudette; BWCoV SW1, Beluga Whale coronavirus SW1; HCoV NL63, human coronavirus NL63; HCoV 229E, human coronavirus 229E; Mi-BatCoV 1B, *Miniopterus* bat coronavirus 1B; Mi-BatCoV HKU8, *Miniopterus* bat coronavirus HKU8; Rh-BatCoV HKU2, *Rhinolophus* bat coronavirus HKU2; Sc-BatCoV 512, *Scotophilus* bat coronavirus 512/05; PEDV CV777, porcine epidemic diarrhoea coronavirus; FCoV VR-2202, feline coronavirus; MHV A59, mouse hepatitis virus A59; HCoV HKU1, human coronavirus HKU1; HCoV OC43, human coronavirus OC43; Ty-BatCoV HKU4, *Tylonycteris* bat coronavirus HKU4; Pi-BatCoV HKU5, *Pipistrellus* bat coronavirus HKU5; Ro-BatCoV HKU9, *Rousettus* bat coronavirus HKU9; SARS-HCoV Fra1, severe acute respiratory syndrome reproductive syndrome virus, strain VR-2332 (referred to as PRRSV in the main text); PRRSV LV4, porcine respiratory and reproductive syndrome virus, strain Lelystad; LDV nv-type-C, lactate dehydrogenase-elevating virus neuro-virulent type C; SHFV LVR-42-0, simian hemorrhagic fever virus; EAV Bucyrus, equine arteritis virus (referred to as EAV in the main text).



**Figure S2.** Phylogenetic tree of the NendoU domain constructed on the basis of the amino acid alignment (174 sites) shown in Figure S1. For virus acronyms, see the legend of figure S1. The scale bar represents 0.1 substitutions per site on average. Numbers at branching points indicate support from 1,000 nonparametric bootstraps.

## REFERENCES

1. **Alger, T. D.** 1970. Copper (II) and manganese (II) effects on ribonuclease A activity. *Biochemistry* **9**:3248-3255
2. **Almazan, F., M. L. Dediego, C. Galan, D. Escors, E. Alvarez, J. Ortego, I. Sola, S. Zuniga, S. Alonso, J. L. Moreno, A. Nogales, C. Capiscol, and L. Enjuanes.** 2006. Construction of a severe acute respiratory syndrome coronavirus infectious cDNA clone and a replicon to study coronavirus RNA synthesis. *J.Virol.* **80**:10900-10906
3. **Bhardwaj, K., L. Guarino, and C. C. Kao.** 2004. The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. *J.Virol.* **78**:12218-12224
4. **Bhardwaj, K., S. Palaninathan, J. M. Alcantara, L. L. Yi, L. Guarino, J. C. Sacchettini, and C. C. Kao.** 2008. Structural and functional analyses of the severe acute respiratory syndrome coronavirus endoribonuclease Nsp15. *J.Biol.Chem.* **283**:3655-3664
5. **Bhardwaj, K., J. Sun, A. Holzenburg, L. A. Guarino, and C. C. Kao.** 2006. RNA recognition and cleavage by the SARS coronavirus endoribonuclease. *J.Mol.Biol.* **361**:243-256
6. **Brierley, I., P. Digard, and S. C. Inglis.** 1989. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* **57**:537-547
7. **Decroly, E., I. Imbert, B. Coutard, M. Bouvet, B. Selisko, K. Alvarez, A. E. Gorbalenya, E. J. Snijder, and B. Canard.** 2008. Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2'O)-methyltransferase activity. *J.Virol.* **82**:8071-8084
8. **Eckerle, L. D., X. Lu, S. M. Sperry, L. Choi, and M. R. Denison.** 2007. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J.Virol.* **81**:12135-12144
9. **Edgar, R. C.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**:1792-1797
10. **Eriksson, K. K., L. Cervantes-Barragan, B. Ludewig, and V. Thiel.** 2008. Mouse Hepatitis Virus liver pathology is dependent on ADP-ribose 1"-phosphatase, a viral function conserved in the alpha-like supergroup. *J.Virol.* **82**: 12325-12334
11. **Gioia, U., P. Laneve, M. Dlakic, M. Arceci, I. Bozzoni, and E. Caffarelli.** 2005. Functional characterization of XendoU, the endoribonuclease involved in small nucleolar RNA biosynthesis. *J.Biol. Chem.* **280**:18996-19002
12. **Gorbalenya, A. E., L. Enjuanes, J. Ziebuhr, and E. J. Snijder.** 2006. Nidovirales: evolving the largest RNA virus genome. *Virus Res.* **117**:17-37
13. **Guarino, L. A., K. Bhardwaj, W. Dong, J. Sun, A. Holzenburg, and C. Kao.** 2005. Mutational analysis of the SARS virus Nsp15 endoribonuclease: identification of residues affecting hexamer formation. *J.Mol.Biol.* **353**:1106-1117
14. **Guindon, S. and O. Gascuel.** 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst.Biol.* **52**:696-704
15. **Imbert, I., J. C. Guillemot, J. M. Bourhis, C. Bussetta, B. Coutard, M. P. Egloff, F. Ferron, A. E. Gorbalenya, and B. Canard.** 2006. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J.* **25**:4933-4942
16. **Ivanov, K. A., T. Hertzog, M. Rozanov, S. Bayer, V. Thiel, A. E. Gorbalenya, and J. Ziebuhr.** 2004. Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc.Natl.Acad. Sci.U.S.A* **101**:12694-12699
17. **Joseph, J. S., K. S. Saikatendu, V. Subramanian, B. W. Neuman, M. J. Buchmeier, R. C. Stevens, and P. Kuhn.** 2007. Crystal structure of a monomeric form of severe acute respiratory syndrome

- coronavirus endonuclease nsp15 suggests a role for hexamerization as an allosteric switch. *J.Virol.* **81**:6700-6708
18. **Kang, H., K. Bhardwaj, Y. Li, S. Palaninathan, J. Sacchettini, L. Guarino, J. L. Leibowitz, and C. C. Kao.** 2007. Biochemical and genetic analyses of murine hepatitis virus Nsp15 endoribonuclease. *J.Virol.* **81**:13587-13597
  19. **Laneve, P., F. Altieri, M. E. Fiori, A. Scalon, I. Bozzoni, and E. Caffarelli.** 2003. Purification, cloning, and characterization of XendoU, a novel endoribonuclease involved in processing of intron-encoded small nucleolar RNAs in *Xenopus laevis*. *J.Biol.Chem.* **278**:13026-13032
  20. **Laneve, P., U. Gioia, R. Ragno, F. Altieri, F. C. Di, T. Santini, M. Arceci, I. Bozzoni, and E. Caffarelli.** 2008. The tumor marker human placental protein 11 is an endoribonuclease. *J.Biol.Chem.* **283**:34712-34719
  21. **Minskaia, E., T. Hertzog, A. E. Gorbalenya, V. Campanacci, C. Cambillau, B. Canard, and J. Ziebuhr.** 2006. Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc.Natl.Acad.Sci.U.S.A* **103**:5108-5113
  22. **Nielsen, H. S., G. Liu, J. Nielsen, M. B. Oleksiewicz, A. Botner, T. Storgaard, and K. S. Faaberg.** 2003. Generation of an infectious clone of VR-2332, a highly virulent North American-type isolate of porcine reproductive and respiratory syndrome virus. *J.Virol.* **77**:3702-3711
  23. **Pasternak, A. O., W. J. Spaan, and E. J. Snijder.** 2006. Nidovirus transcription: how to make sense...? *J.Gen.Virol.* **87**:1403-1421
  24. **Posthuma, C. C., D. D. Nedialkova, J. C. Zevenhoven-Dobbe, J. H. Blokhuis, A. E. Gorbalenya, and E. J. Snijder.** 2006. Site-directed mutagenesis of the Nidovirus replicative endoribonuclease NendoU exerts pleiotropic effects on the arterivirus life cycle. *J.Virol.* **80**:1653-1661
  25. **Putics, A., W. Filipowicz, J. Hall, A. E. Gorbalenya, and J. Ziebuhr.** 2005. ADP-ribose-1"-monophosphatase: a conserved coronavirus enzyme that is dispensable for viral replication in tissue culture. *J.Virol.* **79**:12721-12731
  26. **Raines, R. T.** 1998. Ribonuclease A. *Chem.Rev.* **98**:1045-1066
  27. **Ranjith-Kumar, C. T., L. Gutshall, M. J. Kim, R. T. Sarisky, and C. C. Kao.** 2002. Requirements for de novo initiation of RNA synthesis by recombinant flaviviral RNA-dependent RNA polymerases. *J.Virol.* **76**:12526-12536
  28. **Renzi, F., E. Caffarelli, P. Laneve, I. Bozzoni, M. Brunori, and B. Vallone.** 2006. The structure of the endoribonuclease XendoU: From small nucleolar RNA processing to severe acute respiratory syndrome coronavirus replication. *Proc.Natl.Acad.Sci.U.S.A* **103**:12365-12370
  29. **Ricagno, S., M. P. Egloff, R. Ulferts, B. Coutard, D. Nurizzo, V. Campanacci, C. Cambillau, J. Ziebuhr, and B. Canard.** 2006. Crystal structure and mechanistic determinants of SARS coronavirus nonstructural protein 15 define an endoribonuclease family. *Proc.Natl.Acad.Sci.U.S.A* **103**:11892-11897
  30. **Roth-Cross, J. K., H. Stokes, G. Chang, M. M. Chua, V. Thiel, S. R. Weiss, A. E. Gorbalenya, and S. G. Siddell.** 2009. Organ-Specific Attenuation of Murine Hepatitis Virus (MHV-A59) by Replacement of Catalytic Residues in the Putative Viral Cyclic Phosphodiesterase NS2. *J.Virol.* **83**: 3743-3753
  31. **Sawicki, S. G., D. L. Sawicki, and S. G. Siddell.** 2007. A contemporary view of coronavirus transcription. *J.Virol.* **81**:20-29
  32. **Schutze, H., R. Ulferts, B. Schelle, S. Bayer, H. Granzow, B. Hoffmann, T. C. Mettenleiter, and J. Ziebuhr.** 2006. Characterization of White bream virus reveals a novel genetic cluster of nidoviruses. *J.Virol.* **80**:11598-11609
  33. **Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. M. Poon, Y. Guan, M. Rozanov, W. J. M. Spaan, and A. E. Gorbalenya.** 2003. Unique and conserved features of genome

and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J.Mol. Biol.* **331**:991-1004

34. **Thompson, J. E., F. D. Venegas, and R. T. Raines.** 1994. Energetics of catalysis by ribonucleases: fate of the 2',3'-cyclic phosphodiester intermediate. *Biochemistry* **33**:7408-7414
35. **van Aken, D., J. Zevenhoven-Dobbe, A. E. Gorbalenya, and E. J. Snijder.** 2006. Proteolytic maturation of replicase polyprotein pp1a by the nsp4 main proteinase is essential for equine arteritis virus replication and includes internal cleavage of nsp7. *J.Gen.Virol.* **87**:3473-3482
36. **van den Born, E., A. P. Gultyaev, and E. J. Snijder.** 2004. Secondary structure and function of the 5'-proximal region of the equine arteritis virus RNA genome. *RNA* **10**:424-437
37. **Whelan, S. and N. Goldman.** 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol.Biol.Evol.* **18**:691-699
38. **Xu, X., Y. Zhai, F. Sun, Z. Lou, D. Su, Y. Xu, R. Zhang, A. Joachimiak, X. C. Zhang, M. Bartlam, and Z. Rao.** 2006. New antiviral target revealed by the hexameric structure of mouse hepatitis virus nonstructural protein nsp15. *J.Virol.* **80**:7909-7917
39. **Ziebuhr, J., E. J. Snijder, and A. E. Gorbalenya.** 2000. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J.Gen.Virol.* **81**:853-879

# Chapter 5

## **Arterivirus subgenomic mRNA synthesis and virion biogenesis depend on the multifunctional nsp1 autoprotease**

Marieke A. Tijms\*,  
Danny D. Nedialkova\*,  
Jessica C. Zevenhoven-Dobbe\*,  
Alexander E. Gorbalenya, and  
Eric J. Snijder

\*Equal contribution

J. Virol. 2007. 81(19):10496-505  
Reprinted with permission

## ABSTRACT

Many groups of plus-stranded RNA viruses produce additional, subgenomic mRNAs to regulate the expression of part of their genome. Arteriviruses and coronaviruses (order *Nidovirales*) are unique among plus-stranded RNA viruses for using a mechanism of discontinuous RNA synthesis to produce a nested set of 5'- and 3'-coterminally subgenomic mRNAs, which serve to express the viral structural protein genes. The discontinuous step presumably occurs during minus-strand synthesis and joins noncontiguous sequences copied from the 3'- and 5'-proximal domains of the genomic template. Nidovirus genome amplification ("replication") and subgenomic mRNA synthesis ("transcription") are driven by 13 to 16 nonstructural proteins (nsp's), generated by autocatalytic processing of two large "replicase" polyproteins. Previously, using a replicon system, the N-terminal nsp1 replicase subunit of the arterivirus equine arteritis virus (EAV) was found to be dispensable for replication but crucial for transcription. Using reverse genetics, we have now addressed the role of nsp1 against the background of the complete EAV life cycle. Mutagenesis revealed that nsp1 is in fact a multifunctional regulatory protein. Its papain-like autoprotease domain releases nsp1 from the replicase polyproteins, a cleavage essential for viral RNA synthesis. Several mutations in the putative N-terminal zinc finger domain of nsp1 selectively abolished transcription, while replication was either not affected or even increased. Other nsp1 mutations did not significantly affect either replication or transcription but still dramatically reduced the production of infectious progeny. Thus, nsp1 is involved in at least three consecutive key processes in the EAV life cycle: replicase polyprotein processing, transcription, and virion biogenesis.

abstract



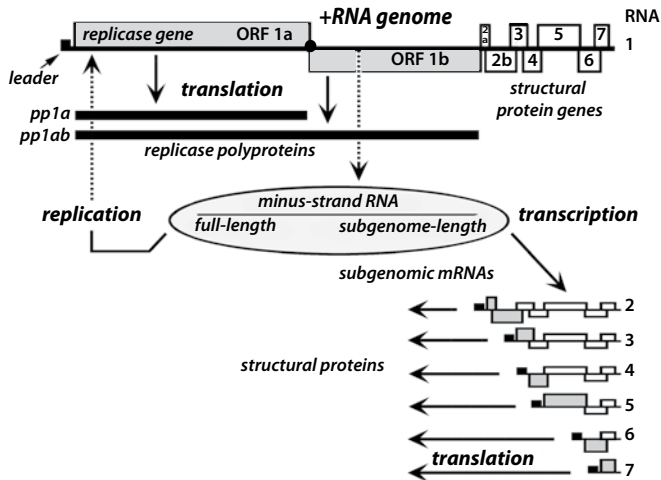
## INTRODUCTION

The replication of plus-stranded RNA viruses of eukaryotes depends on a unique process of cytoplasmic RNA-dependent RNA synthesis that is primarily directed by an enzyme complex, always including an RNA-dependent RNA polymerase (RdRp), that is produced by translation of the incoming viral genome (reviewed, e.g., in references<sup>1, 3, 7</sup>). The “nonstructural proteins” (nsp’s) or “replicase subunits” of mammalian plus-stranded RNA viruses are commonly produced by proteolytic cleavage of larger polyproteins (reviewed e.g., in references<sup>10,14</sup>). This strategy can be supplemented with the synthesis of additional, subgenomic (sg) mRNAs, a process referred to as “transcription” for the purpose of this article, to express additional proteins (reviewed in reference<sup>34</sup>). Members of the order Nidovirales of plus-stranded RNA viruses (*Arteriviridae*, *Coronaviridae*, and *Roniviridae*) rely heavily on the latter strategy (reviewed in references<sup>39, 49,50</sup>) to produce their structural proteins from genes located in the 3′-proximal third of their genome (Fig. 1).

Nidoviruses have been grouped together on the basis of similarities in genome organization and nonstructural and structural protein expression strategies, as well as the presumed immediate common ancestry of key replicative enzymes (see our recent reviews in references<sup>12,56</sup> and references cited therein). The nidovirus replicase/transcriptase, which will be referred to as “replicase” for simplicity, is expressed from the genome RNA by translation of the large open reading frames (ORFs) 1a and 1b (Fig. 1). This yields polyproteins pp1a and pp1ab, with the latter being produced via a conserved ORF1a/ORF1b -1 ribosomal frameshift mechanism (see reference<sup>5</sup> and references cited therein). Nidovirus replicases, in particular those of nidoviruses with large genomes (corona-, toro-, and roniviruses [25 to 31 kb]), include several enzymatic activities that are rare or lacking in other RNA viruses<sup>11, 12, 32, 53, 73</sup>. Triggered by the 2003 outbreak of severe acute respiratory syndrome (SARS)-coronavirus, significant progress has been made in the functional and structural characterization of coronavirus replicative proteins<sup>4, 32, 73</sup>. Still, many questions remain regarding the molecular mechanisms that nidoviruses have evolved to coordinate the various steps of their replicative cycle.

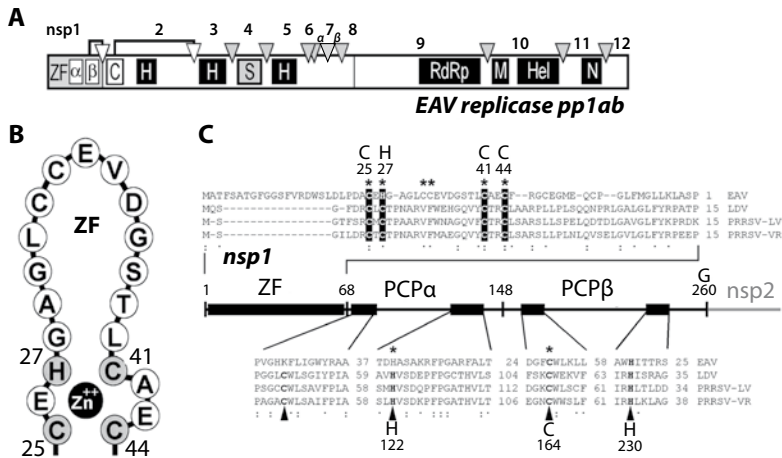
The arterivirus family is the nidovirus branch that stands out for its much smaller genome size (12 to 16 kb). The pp1a and pp1ab replicase polyproteins of the arterivirus prototype equine arteritis virus (EAV)<sup>54</sup> are 1,727 and 3,175 amino acids (aa) long, respectively (for reviews, see references<sup>12, 52, 55</sup>) and are cleaved (Fig. 2A) into 13 nsp’s by viral proteases contained in nsp1, nsp2, and nsp4<sup>63, 75</sup>. Seven structural protein genes, residing in the 3′-terminal quarter of the EAV genome, are expressed from a nested set of sg mRNAs (Fig. 1) and are all dispensable for replication and transcription<sup>36</sup>.

Nidovirus RNA synthesis in infected cells entails both amplification of the viral genome and sg RNA production. The “bodies” of arteri- and coronavirus sg mRNAs are 3′-coterminal, but they also contain a 5′ common leader sequence that is identical to the 5′ end of the genome. Replication and transcription are thought to proceed through different minus-strand intermediates. Whereas a full-length minus-strand template is used for replication, sg mRNAs are assumed to be synthesized from subgenome-length minus



**Figure 1.** Schematic diagram of the genome organization and expression of EAV. The regions encoding the replicase polyproteins (pp1a and pp1ab) and structural proteins are indicated on the genome. The replication/transcription complex, presumably containing the full-length and subgenome-length minus-strand templates for replication and transcription, is depicted below the genome. Subgenomic mRNAs, with the black boxes representing their common 5' leader sequence, are shown in the bottom part of the scheme.

strands, generated by a process of discontinuous RNA synthesis (reviewed in references<sup>39, 48–50, 70</sup>) involving the 3'-terminal extension of the nascent strands (the complements of the sg mRNA bodies) with the complement of the leader sequence. This discontinuous step in minus-strand RNA synthesis, which may resemble copy-choice RNA recombination, is regulated (in part) by conserved transcription-regulating sequences (TRSs). In EAV, both sequence elements and higher-order RNA structures have been postulated to direct the transfer of the nascent minus strand from the body TRS to the leader TRS in the genomic RNA template<sup>38, 40, 41, 64, 65</sup>. In addition to these RNA-RNA interactions, several EAV replicase subunits have been specifically implicated in transcription: the nsp10 helicase<sup>51, 68, 69</sup>, the nsp11 putative endoribonuclease<sup>45</sup>, and nsp1<sup>62</sup>, the most N-terminal replicase cleavage product. Previous studies showed that EAV replication can proceed in the absence of transcription<sup>62, 68</sup>, suggesting that the latter process is in part controlled by transcription-specific replicase functions. Whereas mutations in nsp10 and nsp11 affected both genome amplification and transcription<sup>45, 68</sup>, deletion of the nsp1-coding sequence from the replicase gene selectively blocked sg RNA synthesis<sup>62</sup>, which could be partially restored in a replicon RNA that expressed nsp1 from an artificial alternative locus in the genome (replicon DITRAC)<sup>62</sup>. EAV nsp1 (260 aa) cotranslationally releases itself from the replicase polyproteins through a papain-like cysteine protease (PCP $\beta$ ) activity in its C-terminal half<sup>58, 59</sup>. Comparative sequence analysis identified two other domains in nsp1 (Fig. 2C), an additional PCP (PCP $\alpha$ ), which has become inactivated in the course of EAV evolution but is functional in other arteriviruses<sup>9</sup>, and a potential N-terminal zinc finger (ZF) domain (Fig. 2B)<sup>62</sup>. In the artificial setting of the DITRAC replicon, the ZF domain and



**Figure 2.** (A) Processing scheme of the EAV replicase, depicted in the form of the pp1ab polyprotein. The three viral protease domains (the nsp1 PCPβ [β], the nsp2 cysteine proteinase [CP], and the nsp4 serine proteinase [S]) and their cleavage sites are indicated. Other domain abbreviations are as follows: ZF, nsp1 zinc finger domain; α, nsp1 PCPα domain; H, hydrophobic domains; RdRp, RNA-dependent RNA polymerase (nsp9); M, metal-binding domain (nsp10); Hel, helicase (nsp10); N, nidovirus-specific endoribonuclease NendoU (nsp11). (B) Schematic representation of the putative nsp1 ZF domain, showing the four residues that are proposed to coordinate zinc in gray. (C) Comparison of key sequences of the three domains identified in the arterivirus nsp1 region, ZF, PCPα, and PCPβ. The putative zinc-coordinating residues of the ZF domain and the active-site Cys and His residues of the two PCP domains are indicated. Note that EAV PCPα is no longer active due to the loss of its active-site Cys (9). Residues targeted by mutagenesis in this study are indicated with an asterisk.

the structural integrity of nsp1, but not the proteolytic activity of PCPβ, were important for transcription<sup>62</sup>. The role of nsp1 has now been addressed in the context of the full-length EAV genome and the complete viral life cycle. We conclude that EAV nsp1 not only couples replicase maturation to transcription but also plays a surprising, yet unexplained, role in virion biogenesis. Several nsp1 mutations that did not dramatically affect viral RNA synthesis were found to severely reduce the production of infectious progeny. Thus, EAV nsp1 is a multifunctional protein that is involved in at least three consecutive key processes in the EAV life cycle: proteolytic maturation; of the replicase, transcription, and virion biogenesis.

## MATERIALS AND METHODS

### EAV reverse genetics

Site-directed PCR mutagenesis<sup>25</sup> was used to engineer mutations in shuttle plasmids carrying the nsp1 coding sequence. Following sequence analysis, mutations were transferred to EAV full-length cDNA clone pEAV211<sup>64, 68</sup>. The mutants of the PCPβ active-site cysteine

(Cys-164) and the nsp1/2 cleavage site (Gly-260/Gly-261) have been described before<sup>58</sup>. The amino acid substitutions engineered to generate novel nsp1 mutants are listed in Table 1.

Following *in vitro* transcription of full-length cDNA clones by T7 RNA polymerase, BHK-21 cells were transfected with equal amounts of full-length EAV RNA as described previously<sup>68</sup>. Immunofluorescence assays (IFAs)<sup>67,68</sup> were used to monitor transfection efficiencies.

### **Analysis of virus mutants**

For an initial assessment of viral phenotypes (in terms of replication and transcription), dual-labeling IFAs<sup>67</sup> were performed with a rabbit antiserum specific for nsp3<sup>43</sup>, a replicase cleavage product expressed from the genome, and a mouse monoclonal antibody (MAb) recognizing the nucleocapsid (N) protein (N MAb 3E2)<sup>29</sup>, which is expressed from the smallest sg mRNA. Viral RNA synthesis was studied in more detail by isolating total intracellular RNA at the end of the first replication cycle (~14 h post-transfection [hpt]) using the acidic phenol extraction method, as described previously<sup>70</sup>. RNA was separated in denaturing formaldehyde-agarose gels, and viral mRNAs were visualized by hybridization with a radioactively labeled, antisense oligonucleotide probe (complementary to the 3' end of the genome), which recognizes both genomic and all sg mRNAs<sup>70</sup>. Supernatants harvested from transfected cell cultures were analyzed for infectious progeny by plaque assays<sup>36</sup> and infection of fresh BHK-21 cells.

### **PCP $\beta$ activity assay**

A vector (pEAV $\Delta$ H) for testing PCP $\beta$  activity in an *in vitro* transcription and translation assay<sup>58</sup> was engineered by making an internal HindIII deletion (removing nucleotides [nt] 1506 to 12308 of the EAV genome; NCBI accession no. NC\_002532) in the full-length cDNA clones for selected mutants. These constructs contained an in-frame fusion of the 5' end of ORF1a to the 3' end of ORF6 and encoded a product of 453 aa (52 kDa) including the full-length nsp1 and the nsp1/2 cleavage site. The fusion gene was preceded by a T7 promoter and the natural genomic 5'-untranslated region and followed by the genomic 3'-untranslated region and a poly(A) tail. The *in vitro* PCP assay was performed by translation of transcripts derived from these vectors in a rabbit reticulocyte lysate (Promega). The reaction was performed at 30°C for 90 min in the presence of [<sup>35</sup>S]methionine, after which translation and cleavage products were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and autoradiography.

### **Sequence analysis and RNA structure prediction**

The RNA structure of the 5'-proximal region of EAV ORF1a was predicted using the genetic algorithm of STAR v4.4<sup>19</sup> and the Zuker algorithm<sup>77</sup> on the M-fold web server<sup>78</sup>. Sequence alignments were created using the ClustalW algorithm incorporated in Vector NTI (Invitrogen). The hairpin structure predicted to underlie the ZF-coding region (see

Fig. 6) was mutated in full-length cDNA clone pEAV211-RNAko using a combination of seven translationally silent mutations in codons specifying residues 22 to 28 of nsp1 (Fig. 6): C-290→A, C-293→U, U-296→A,U-299→C, G-302→A, U-305→C, and C-308→G (numbers refer to positions in the EAV genome sequence).

## RESULTS

### Design and characterization of EAV nsp1 mutants

We have previously reported that nsp1 is dispensable for EAV genome replication but absolutely required for transcription<sup>62</sup>. In that study, expression of an engineered wild-type nsp1 cassette from an artificial internal ribosomal entry site (IRES) inserted in the 3'-terminal part of replicon DITRAC could complement for the deletion of the natural nsp1-coding sequence from the replicase gene<sup>62</sup>. This resulted in partial restoration of transcription, but since the IRES-nsp1 cassette in the DITRAC replicon replaced five of the seven structural protein genes, virus production could not be studied in this system. To address the role of nsp1 against the background of the complete EAV life cycle, we have now used a reverse-genetics approach based on the use of full-length cDNA clones<sup>68</sup>. Full-length RNA derived from cDNA clones containing nsp1 mutations (Table 1) was transfected into BHK-21 cells. Replication and transcription were initially monitored using a previously described dual-labeling IFA<sup>62</sup> detecting nonstructural and structural protein expression (data not shown). Subsequently, additional assays were used to characterize the phenotypes. The production of infectious progeny was investigated by monitoring the spread of the infection in transfected cell cultures and/or using cell culture supernatants to infect fresh cells.

### The nsp1/2 cleavage is essential for EAV genome replication

We first analyzed the importance of cleavage of the nsp1/2 site by the C-terminal PCPβ domain of nsp1<sup>58</sup>. In the setting of the DITRAC replicon, PCPβ activity was dispensable for the functionality of nsp1 in transcription<sup>62</sup>. However, in DITRAC, nsp1 was expressed from a separate IRES-driven cistron and did not have to play its usual role in replicase proteolysis. In this replicon, cleavage of the nsp1/2 bond was effectively substituted for by translation initiation at the 5' end of the nsp2 coding sequence and independent expression of nsp1 from the 3'-proximal part of the genome.

Using a full-length cDNA clone, cleavage of the nsp1/2 site was now found to be a prerequisite for EAV RNA synthesis. Several mutations that were previously reported to abolish processing of the nsp1/2 site<sup>58,59</sup> rendered the viral RNA noninfectious. These included replacement of the PCPβ active-site Cys-164 with Ser and replacement of the P1 (Gly-260) or P1' (Gly-261) residues of the nsp1/2 cleavage site with Val (Table 1) (data not shown). The lethal effect of these mutations was evident, in several independent experiments, from the complete lack of IFA signal in transfected cells and the absence of viral RNA synthesis and infectious progeny.

**Table 1.** Overview of the genotype and first-cycle phenotype of EAV nsp1 mutants used in this study

Construct	Mutation	Codon <sup>a</sup>		Replication <sup>b</sup>	Transcription <sup>b</sup>	Plaque size	Titer (PFU/ml) <sup>c</sup>	Phenotype summary <sup>d</sup>
		Wild type	Mutant					
pEAV211	NA <sup>e</sup>	NA	NA	Normal	Normal	Normal	10 <sup>7</sup> -10 <sup>8</sup>	Wild-type control
C25A	Cys -25→Ala	UGU	<b>GCU</b>	Enhanced	Negative	No plaques	<2x10 <sup>1</sup>	Transcription negative; nonviable
C25H	Cys -25→His	UGU	<b>CAU</b>	Normal	Normal	Small	10 <sup>5</sup> -10 <sup>6</sup>	Transcription ~normal; attenuated
H27A	His-27→Ala	CAU	<b>GCC</b>	Normal	Normal	Small	10 <sup>2</sup> -10 <sup>3</sup>	Transcription ~normal; attenuated
H27C	His-27→Cys	CAU	<b>UGU</b>	Normal	Normal	Small	10 <sup>2</sup> -10 <sup>3</sup>	Transcription ~normal; attenuated
H27R	His-27→Arg	CAU	<b>CGU</b>	Enhanced	Reduced	Normal	10 <sup>2</sup> -10 <sup>3</sup>	Transcription impaired; rapid reversion
C32A	Cys -32→Ala	UGC	<b>GCG</b>	Normal	Normal	Normal	10 <sup>7</sup> -10 <sup>8</sup>	Wild type
C33A	Cys -33→Ala	UGC	<b>GCC</b>					
C41A	Cys -41→Ala	UGC	<b>GCG</b>	Enhanced	Negative	No plaques	<2x10 <sup>1</sup>	Transcription negative; nonviable
C41H	Cys -41→His	UGC	<b>CAU</b>	Enhanced	Negative	No plaques	<2x10 <sup>1</sup>	Transcription negative; nonviable
C44A	Cys -44→Ala	UGU	<b>GCU</b>	Enhanced	Negative	No plaques	<2x10 <sup>1</sup>	Transcription negative; nonviable
C44H	Cys -44→His	UGU	<b>CAU</b>	Enhanced	Negative	No plaques	<2x10 <sup>1</sup>	Transcription negative; nonviable
C164S	Cys-164→Ser	UGC	<b>AGC</b>	Negative	Negative	No plaques	<2x10 <sup>1</sup>	Replication negative; nonviable
G260V	Gly-260→Val	GGC	<b>GUA</b>	Negative	Negative	No plaques	<2x10 <sup>1</sup>	Replication negative; nonviable
G261V	Gly-261→Val	GGC	<b>GUA</b>	Negative	Negative	No plaques	<2x10 <sup>1</sup>	Replication negative; nonviable
RNAko	7 silent mutations	See text	See text	Normal	Normal	Normal	10 <sup>7</sup> -10 <sup>8</sup>	Wild type

<sup>a</sup> Nucleotide substitutions are indicated in boldface.

<sup>b</sup> Transfected cells were analyzed by IFA and Northern blot analysis at the end of the first cycle of replication (14 to 16 hpt).

<sup>c</sup> Infectious progeny titers in plaque assays of medium from transfected cells harvested at 18 hpt.

<sup>d</sup> Depending on the mutant, data from two to four independent transfection experiments are summarized.

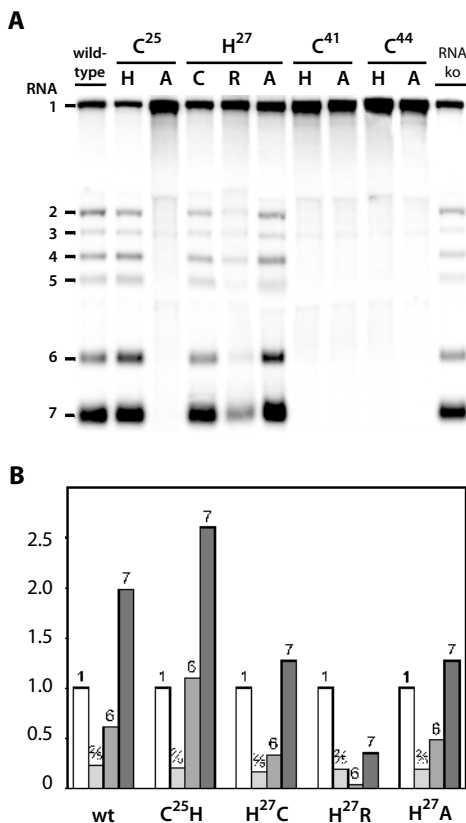
<sup>e</sup> NA, not applicable.

These results are in contrast with our previous observation that deletion of nsp1 selectively blocks transcription but not replication<sup>62</sup>. We therefore conclude that cleavage of the nsp1/2 site by the nsp1 PCPβ domain must be essential to produce a functional EAV replication complex. Possibly, the N terminus of nsp2 has to be liberated for this subunit to fulfill its role in the formation of the membrane-bound replication complex<sup>43, 57</sup>. This

interpretation is further supported by the nonviable phenotype of an EAV construct, in which green fluorescent protein (GFP) was fused to the N terminus of nsp2<sup>66</sup>.

### The nsp1 ZF domain plays a critical role in transcription

On the basis of site-directed mutagenesis of Cys-25 and Cys-44 (both residues changed to Ala) in the DITRAC replicon, the nsp1 ZF domain was previously implicated in regulation of transcription<sup>62</sup>. In the context of the EAV full-length clone, the ZF domain was now probed more extensively (Table 1). The four putative zinc-coordinating residues (Cys-25, His-27, Cys-41, and Cys-44) were targeted, and - as a negative control - two Cys residues (Cys-32 and Cys-33) that are not conserved in arterivirus nsp1 sequences (Fig. 2C) were also replaced. Nonconservative mutagenesis (to Ala) of the latter two residues resulted in a wild-type phenotype, indicating that Cys residues at positions 32 and 33 of the EAV ZF domain are dispensable for replication in cell culture. The four putative zinc-coordinating residues were replaced both nonconservatively (Cys or His to Ala or Arg) and conservatively (Cys to His and His to Cys). The latter strategy was aimed at maintaining (partial)



**Figure 3.** (A) Northern blot analysis of genome (RNA1) and sg mRNA (RNAs 2 to 7) synthesis in BHK-21 cells transfected with EAV nsp1 ZF mutants (14 hpt) and the RNA<sub>ko</sub> mutant designed to probe the role of the RNA structure of the ZF-coding region (see text). RNA was isolated at 14 hpt, and hybridization was carried out using a <sup>32</sup>P-labeled oligonucleotide probe complementary to the 3' end of the genome and thus recognizing all viral mRNA species. Note the upregulation of genome synthesis in transcription-negative mutants. (B) Phosphorimager analysis of the ratio of sg mRNA to genome accumulation, based on the gel depicted in panel A. For the transcription-positive mutants, the amount of genome was put at 1 and the relative amounts measured for mRNA7, mRNA6, and mRNAs 2 to 5 are shown. With the exception of the His-27→Arg mutant, all mutants were concluded to maintain fairly normal transcription levels. wt, wild type.

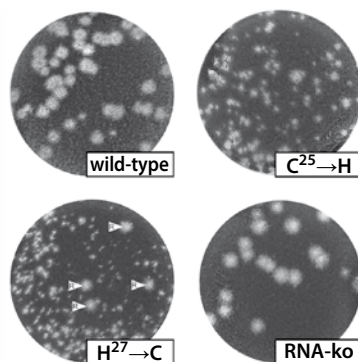
zinc-binding activity. Interestingly, a differential effect of the Ala/Arg versus His/Cys ZF mutations was indeed observed at some positions (Table 1).

Residues Cys-41 and Cys-44 were both absolutely required for the function of nsp1 in transcription (Fig. 3 and Table 1), since their replacement essentially exerted the same effect as deletion of nsp1: transcription was reduced to a level that could not even be detected by using reverse transcription-PCR (RT-PCR) (data not shown). Remarkably, these transcription-negative mutants also displayed a clear upregulation (2.5- to 3-fold) of genome replication (Fig. 3A), suggesting that EAV replication and transcription normally compete for common factors.

### Multiple nsp1 ZF domain mutants are transcription positive but severely impaired in virus production

In terms of transcription, the role of Cys-25 was found to be somewhat less critical than those of Cys-41 and Cys-44. Northern blot analysis revealed that the Cys-25→His mutant produced approximately normal amounts of sg mRNAs (Fig. 3B), although replacement with Ala completely eliminated transcription. Immunoprecipitation analyses and IFAs (D. D. Nedialkova et al., unpublished data) confirmed the synthesis of several structural proteins. Remarkably, however, progeny titers were 100- to 1,000-fold reduced and mutant C25H displayed a small-plaque phenotype (Fig. 4).

Replacement of the fourth putative zinc-coordinating residue, His-27, with either Ala or Cys also yielded a transcription-positive phenotype (Fig. 3). sg RNA synthesis appeared somewhat reduced relative to genome synthesis but not more than twofold (Fig. 3B). However, the drop in infectivity for these two His-27 mutants was more pronounced (4 to 5 logs). Signs of probable reversion were occasionally observed in the form of larger plaques (Fig. 4). The His-27→Arg mutant, on the other hand, displayed a clear transcrip-



**Figure 4.** Plaque phenotype of selected mutants from this study. Virus was harvested from transfected BHK-21 cells at 18 hpt. Plaque assays were also performed on BHK-21 cells and fixed after 3 days. The small-plaque phenotype of the Cys-25→His and His-27→Cys mutants is illustrated, with the latter showing sign of probable rapid reversion (plaques marked R). The fourth photograph illustrates the wild-type plaque phenotype of the RNAko mutant.



tion defect (Fig. 3A and B). Some infectious progeny were produced, but in view of their wild-type size, these plaques likely resulted from rapid reversion of the single-nucleotide mutation introduced to create the H27R mutant. All other mutants in our data set required two replacements to revert to the wild-type codon (Table 1), and they were, accordingly, relatively stable.

The characterization of the transcription-positive nsp1 mutants was extended to several other time points after transfection (between 11 and 18 hpt) with essentially identical results, indicating that these mutants are not merely delayed in terms of virus production but fail to produce the normal level of infectious progeny—an observation implicating nsp1 in an additional step of the viral life cycle, downstream of transcription. Analysis of this phenotype in a longer-time-course experiment was complicated by the emergence and spread of revertants, an issue that will be addressed in a follow-up study.

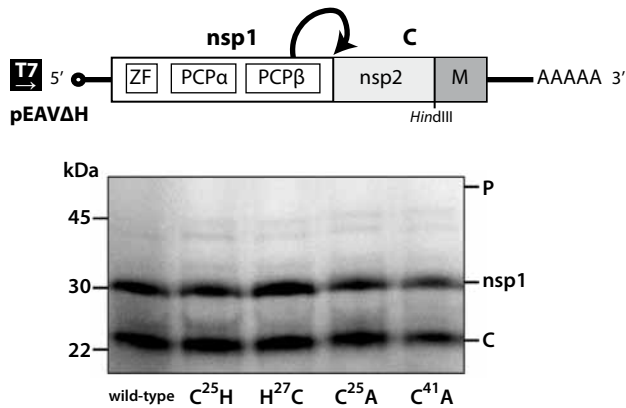
### **PCP $\beta$ proteolytic activity is not affected by mutations in the ZF domain**

In view of the critical importance of cleavage at the nsp1/2 site for virus viability (see above), we selected a number of ZF mutants to investigate whether their nsp1 PCP $\beta$  activity might have been compromised. PCP $\beta$  was previously shown to be highly efficient in directing the nsp1/2 cleavage in a rabbit reticulocyte lysate-based *in vitro* translation system<sup>58</sup>. By making an internal deletion in the full-length cDNA clones for these mutants, we engineered a vector (pEAV $\Delta$ H) for *in vitro* transcription and translation. The pEAV $\Delta$ H product was a 52-kDa (453-aa) fusion protein consisting of the full-length nsp1, the N-terminal part of nsp2, and the C-terminal part of the M protein, which is encoded by ORF6.

*In vitro* transcripts derived from these vectors were translated in a reticulocyte lysate in the presence of [<sup>35</sup>S]methionine. Reaction products were analyzed directly by SDS-PAGE and autoradiography (Fig. 5). Wild-type nsp1 showed complete cleavage of the 52-kDa precursor into nsp1 (30 kDa) and the C-terminal 22-kDa fragment. The same observation was made for nsp1 ZF mutants showing (C25H and H27C) or completely lacking (C25A and C41A) sg RNA synthesis, strongly suggesting that the transcription defect in the latter mutants is unrelated to a change in PCP $\beta$  proteolytic activity.

### **Analysis of RNA structure of the nsp1 ZF-coding region**

The 5'-proximal domain of the arterivirus genome is a multifunctional region containing RNA signals that are known to be involved in translation, replication, and transcription. These signals were previously found to extend into the coding part of the genome, more specifically the 5' end of ORF1a<sup>64, 65</sup>. This raised the possibility that defects observed in our nsp1 ZF mutants could be due to the effect of the introduced mutations on RNA sequence or structure, rather than on the nsp1 amino acid sequence. In particular, such RNA-specific effects might explain transcription defects, in view of the proximity of the leader TRS and leader TRS hairpin<sup>64, 65</sup>. Also, the production of progeny virus might be

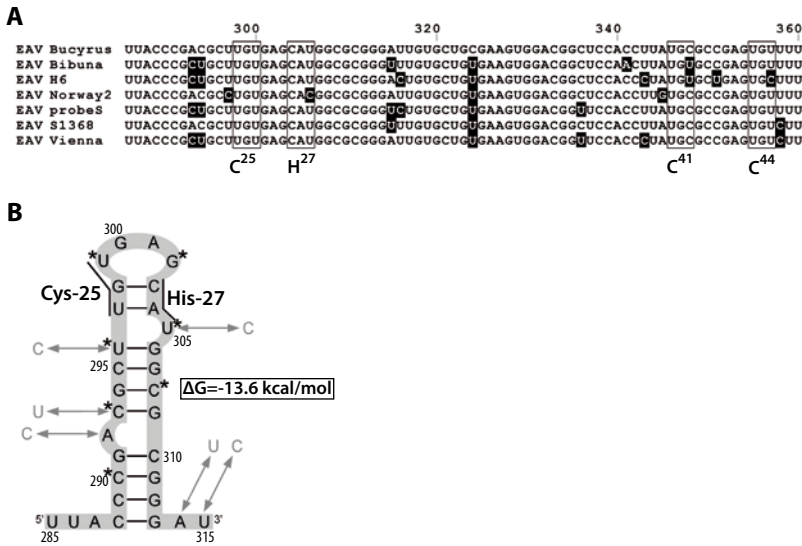


**Figure 5.** *In vitro* assay for PCP $\beta$  activity in nsp1 ZF mutants. Expression construct pEAV $\Delta$ H (see Materials and Methods) was used to *in vitro* translate an nsp1-containing fusion protein in the presence of [ $^{35}$ S]methionine. Reaction products were analyzed by SDS-PAGE and autoradiography. Cleavage of the 52-kDa precursor (P) yields nsp1 and a 22-kDa C-terminal fragment (C). All nsp1 ZF mutants tested showed wild-type nsp1 PCP $\beta$  activity, thus ruling out that a defect in nsp1/2 cleavage was the basis for their phenotype.

affected at the level of genome packaging. The location of encapsidation signal(s) in the EAV genome is still unknown, but previous studies with defective interfering RNAs implicated sequences in the first ~600 nt of the genome in RNA packaging<sup>35</sup>.

A comparative analysis of the ZF-coding sequence (nt 285 to 360) of a number of EAV isolates revealed that this area is highly conserved (Fig. 6A). RNA structure predictions using different methods (see Materials and Methods) revealed a prominent and potentially stable hairpin structure (predicted  $\Delta G$  using Mfold is -13.6 kcal/mol [Fig. 6B]), which can be formed by nt 285 to 315 of the EAV genome. These nucleotides specify nsp1 aa 22 to 30, thus including key ZF residues Cys-25 and His-27, for which mutagenesis had yielded remarkably variable results (see above). Strikingly, the comparative sequence analysis in Fig. 6A provided some additional support for this structure, since four variations found in this region were predicted either to maintain the predicted structure (Fig. 6B; C-293 $\rightarrow$ U and U296 $\rightarrow$ C), to not affect it due to their position in a bulge (A-292 $\rightarrow$ C and U305 $\rightarrow$ C), or to immediately flank the predicted hairpin (A-314 $\rightarrow$ U and U315 $\rightarrow$ C).

In order to address the question whether the defects observed for (some) nsp1 mutants might be due to an effect on the predicted hairpin rather than the nsp1 protein, we engineered a mutant (RNAko) in which a combination of seven translationally silent mutations was introduced in the codons specifying nsp1 residues 22 to 28 (Fig. 6B). Mfold analysis indicated that this combination of mutations in the region of nt 290 to 308 should destroy the predicted hairpin (data not shown). Also some alternative RNA structures that can be predicted for the wild-type sequence would be significantly altered by this combination of mutations. Nevertheless, in multiple experiments, the phenotype of the RNAko mutant was found to be identical to that of wild-type virus, showing normal levels of replication and transcription (Fig. 3), wild-type virus titers (Table 1), and a normal



**Figure 6.** (A) Sequence alignment of the nsp1 ZF-coding domain for a selection of EAV isolates (for details, see reference<sup>64</sup>). The codons for the four key residues of the nsp1 ZF are boxed. (B) Structure of a predicted RNA hairpin that can be formed by the sequence specifying nsp1 ZF residues 22 to 28, thus including the highlighted codons for Cys-25 and His-27. Natural sequence variations (see panel A) are indicated in gray. Residues targeted for mutagenesis in the RNAko mutant (see text) are indicated with asterisks.

plaque size (Fig. 4). Thus, regardless whether the hairpin depicted in Fig. 6 exists or not, it is irrelevant in the context of the results obtained with the Cys-25 and His-27 nsp1 ZF mutants. In this light, also the phenotype of the H27R mutant is important, since its A-304→G mutation is predicted to maintain the RNA hairpin, exchanging a U-A for a U-G base pair. Still, H27R transcription is seriously affected (Fig. 3), again suggesting that it is the nsp1 amino acid sequence that is important rather than the RNA structure of this region.

## DISCUSSION

Following uncoating, plus-strand RNA virus genomes are first translated by host ribosomes to produce the enzymes required for viral RNA synthesis. nsp1 of EAV is the first viral protein expressed and thus seems to occupy a strategic position in the viral life cycle, which may explain the multifunctionality of this replicase subunit that is now emerging. In addition to its critical role in replicase polyprotein processing and sg RNA production described before<sup>62</sup> and extended here, our new data implicate nsp1 in an as-yet-undefined function downstream of RNA synthesis that is ultimately important for virion biogenesis.

The data presented in this article, in particular those obtained from mutagenesis of the putative zinc-binding residues of the ZF domain, confirm the crucial role of nsp1 in

5

EAV transcription. However, whereas EAV proved to be highly sensitive to replacement of Cys-41 and Cys-44, the Cys-25→His substitution was tolerated and the constraints on the His occupying position 27 appeared to be even more relaxed. The fact that - compared to the distal pair of Zn-binding residues (Cys-41 and Cys-44) - mutations at the proximal Cys-25/His-27 pair were relatively well tolerated correlates with (and supports) the alignment presented in Fig. 2C, which shows that the position of His-27 is occupied by Cys in other arteriviruses.

140

Zinc binding by the nsp1 ZF domain remains to be verified, an issue currently pursued by purification and characterization of recombinant nsp1 from *Escherichia coli*. Also, we cannot formally exclude the theoretical possibility that this region of nsp1 contains, e.g., two small adjacent (or even overlapping) domains with different functions. However, the results obtained upon mutagenesis of the four putative zinc-coordinating residues lend credibility to the hypothesis that they are part of a single ZF domain: e.g., mutations in either pair of proposed zinc-binding residues can produce the same transcription-negative phenotype (compare C25A with the C41/C44 mutants [Fig. 3]). The three transcription-positive ZF mutants (C25H, H27C, and H27A) were also the basis for a second important observation, which was that nsp1 mutations can result in a major defect in the production of infectious progeny, thus implicating the protein in an additional step of the viral life cycle, downstream of transcription. The wild-type phenotype of the RNAko mutant (Fig. 6) makes it highly unlikely that disruption of RNA structures in the nsp1 ZF-coding region played a role in the mixed phenotypes that were observed.

Zinc finger domains are short, independently folded stretches of amino acids that require the coordination of one or more zinc atoms to stabilize their higher-order structure<sup>24</sup>. These motifs were first identified in TFIIIA of *Xenopus laevis* and were implicated in nucleic acid recognition<sup>33</sup>. Since then, ZF domains have emerged as one of the most prevalent motifs in the mammalian proteome. They are frequently classified according to the sequential nature of the zinc-coordinating residues, with the CCHH class being the most common one. ZF domains are widely recognized as sequence-specific DNA-binding motifs, but they have also been shown to bind RNA<sup>28</sup>, as well as to participate in protein-protein interactions<sup>27,30</sup>, including dimerization<sup>31</sup>. Given the versatility and diversity of this structural motif, it is not surprising that many viruses also encode ZF proteins. One of the best-characterized examples is the nucleocapsid (NC) protein of onco- and lentiviruses, which contains one or two copies of a CCHC zinc-coordinating domain that is essential for RNA packaging during virus assembly<sup>8,31</sup>. Upon conservative replacement of zinc-binding residues, zinc coordination is maintained and viral genome packaging is not affected, but the resulting mutant viruses are almost completely noninfectious<sup>15,16</sup>. Further analysis implicated these ZF domains in viral RNA synthesis and RNA-protein interactions, thus revealing additional, previously unknown functions of these domains<sup>16,20,26</sup>. Interestingly, a structural study of zinc fingers of the CCHC type in lentiviruses showed that an HCHC mutant was no longer able to bind zinc tetrahedrally, while the CCHH form could still bind zinc tightly, although this resulted in important structural modifications that could be responsible for the loss of infectivity observed for this mutant<sup>47</sup>.

Despite clear differences in protein function and origin, intriguing parallels can be drawn between these retrovirus NC data and our analysis of EAV nsp1 mutants in which putative zinc-coordinating residues were replaced. The transcription-negative phenotype of the Cys-25→A and Cys-41/Cys-44 mutants suggests a complete loss of zinc coordination. However, the conservative replacements of Cys-25→His and His-27→Cys, which in theory may leave the zinc-coordinating potential intact, were both largely tolerated in terms of sg RNA production (Fig. 3), although these mutants were significantly impaired in a later stage of the viral life cycle, possibly virion biogenesis. The His-27→Ala replacement might represent an intermediate that is still able to bind zinc, as described for the distal ZF of the HIV-1 NC (61), in which the same mutant was concluded to retain the tetrahedral coordination of the metal ion, with the vacant ligand position presumably occupied by a water molecule. A similar situation has been described for the CCCH ZF domain in another viral protease, NS3 of hepatitis C virus<sup>22</sup>.

Very little is currently known about the protein-protein and protein-RNA interactions that regulate arterivirus sg RNA synthesis. nsp1 might be involved in the stalling of the viral RdRp complex at the body TRSs by interacting with this conserved RNA sequence. It could also participate in the targeting of the nascent subgenome-length minus strand to the leader TRS in the 5' end of the template or may facilitate the base pairing between body TRS complement and leader TRS - steps that are presumed essential for arterivirus sg RNA synthesis. Regardless of the molecular basis of nsp1 involvement in EAV transcription, the integrity of the ZF domain was also found to be essential for virus production.

Obviously, in the arterivirus system (Fig. 1), structural protein synthesis and consequently virus production are directly controlled by sg mRNA synthesis. It should also be noted that EAV assembly involves a total of seven structural proteins, which are all essential for the production of infectious progeny<sup>36, 71, 72</sup>, and that little is known about the molecular interactions between these components and the kinetics of virus assembly. In other words, a small disturbance at the level of sg mRNA synthesis (e.g., affecting the ratio at which some of the structural proteins are produced) could in theory have major consequences for virus production. Still, three of our mutants (C25H, H27C, and H27A) coupled a deviation of not more than twofold in the ratio of genome to sg mRNA synthesis to a 2- to 5-log drop in infectious progeny titers (Table 1). We consider it highly unlikely, but not impossible, that this is solely due to relatively minor changes at the level of sg mRNA production. On the other hand, ostensibly supporting a direct role of nsp1 in virion biogenesis, there are precedents both for the presence of replicase proteins of plus-strand RNA viruses in virions<sup>37, 42, 46</sup> and for a link between mutations in replicase subunits and the production of infectious progeny<sup>23</sup>.

A role of nsp1 in translation of viral RNA also needs to be investigated, although a preliminary analysis did not reveal major deviations in structural protein synthesis for mutants like C25H (D. D. Nedialkova et al., unpublished data). As in the case of the function of nsp1 in transcription, a better understanding of this phenotype and the role of nsp1 will require identification of interaction partners, which may be either RNA sequences or

other viral proteins. Such studies are currently ongoing using biochemical methods and attempting to isolate second-site revertants for selected nsp1 mutants.

Finally, the data in this paper underlines the importance of the nsp1 PCP $\beta$  domain in the EAV system. This enzyme is a representative of a class of papain-like “accessory proteases” (a term used to discriminate them from the viral “main” protease), which are found in evolutionarily distinct lineages of plus strand RNA viruses<sup>12, 13, 44</sup>. Arteriviruses contain an array of two to four of these enzymes in the N-terminal domain of their replicase (9, 58, 60). Distant homologs of these arterivirus domains are also found in the large nsp3 product of coronaviruses<sup>2, 21, 76</sup>, and it has been postulated that, in the course of nidovirus evolution, these N-terminal papain-like protease domains may have been optimally suited for acquisition and integration of new functions into an expanding replicase<sup>12</sup>. In the case of coronaviruses, engineered deletions and mutations affecting the nsp1 to -3 region and its proteolytic processing by nsp3 accessory proteinases can induce a variety of effects, but, remarkably, a large number of such mutations are tolerated to a certain extent<sup>6, 17, 18, 74</sup>. Thus far, a strictly transcription-specific function, as documented here for EAV nsp1, has not been identified in the coronavirus system, suggesting that arterivirus evolution may have taken a unique turn, by first acquiring multiple accessory PCP-containing replicase subunits and subsequently evolving a key regulatory position for these proteins in the viral life cycle.

## ACKNOWLEDGEMENTS

We are grateful to Erwin van den Born, Martijn van Hemert, Richard Molenkamp, Alexander Pasternak, Clara Posthuma, and Willy Spaan for helpful comments and suggestions. We thank Erwin van den Born and Nancy Beerens for assistance with RNA structure predictions.

M.A.T. and D.D.N. were supported by grants 348-003 and 700.52.306, respectively, from the Council for Chemical Sciences of The Netherlands Organization for Scientific Research (NWO-CW).

## REFERENCES

1. **Ahluquist, P.** 2006. Parallels among positive-strand RNA viruses, reversetranscribing viruses and double-stranded RNA viruses. *Nat. Rev. Microbiol.* **4**:371–382.
2. **Baker, S. C., N. La Monica, C. K. Shieh, and M. M. C. Lai.** 1990. Murine coronavirus gene 1 polyprotein contains an autoproteolytic activity. *Adv. Exp. Biol. Med.* **276**:283–289.
3. **Ball, L. A.** 2001. Replication strategies of RNA viruses, 105–118. *In* D. M. Knipe and P. M. Howley (ed.), *Fields virology*. Lippincott, Williams & Wilkins, Philadelphia, PA.
4. **Bartlam, M., H. Yang, and Z. Rao.** 2005. Structural insights into SARS-coronavirus proteins. *Curr. Opin. Struct. Biol.* **6**:664–672.
5. **Brierley, I., and F. J. Dos Ramos.** 2006. Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res.* **119**:29–42.
6. **Brockway, S. M., and M. R. Denison.** 2005. Mutagenesis of the murine hepatitis virus nsp1-coding region identifies residues important for protein processing, viral RNA synthesis, and viral replication. *Virology* **340**:209–223.
7. **Buck, K. W.** 1996. Comparison of the replication of positive-stranded RNA viruses of plants and animals. *Adv. Virus Res.* **47**:159–251.
8. **Dannull, J., A. Surovoy, G. Jung, and K. Moelling.** 1994. Specific binding of HIV-1 nucleocapsid protein to PSI RNA *in vitro* requires N-terminal zinc finger and flanking basic amino acid residues. *EMBO J.* **13**:1525–1533.
9. **den Boon, J. A., K. S. Faaberg, J. J. M. Meulenberg, A. L. M. Wassenaar, P. G. W. Plagemann, A. E. Gorbalenya, and E. J. Snijder.** 1995. Processing and evolution of the N-terminal region of the arterivirus replicase ORF1a protein: identification of two papainlike cysteine proteases. *J. Virol.* **69**:4500–4505.
10. **Dougherty, W. G., and B. L. Semler.** 1993. Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. *Microbiol. Rev.* **57**:781–822.
11. **Gorbalenya, A. E.** 2001. Big nidovirus genome: when count and order of domains matter. *Adv. Exp. Biol. Med.* **494**:1–17.
12. **Gorbalenya, A. E., L. Enjuanes, J. Ziebuhr, and E. J. Snijder.** 2006. Nidovirales: evolving the largest RNA virus genome. *Virus Res.* **117**:17–37.
13. **Gorbalenya, A. E., E. V. Koonin, and M. M. C. Lai.** 1991. Putative papain-related thiol proteases of positive-strand RNA viruses. Identification of rubiand aphthovirus proteases and delineation of a novel conserved domain associated with proteases of rubi-, alpha- and coronaviruses. *FEBS Lett.* **288**:201–205.
14. **Gorbalenya, A. E., and E. J. Snijder.** 1996. Viral cysteine proteases. *Persp. Drug Discov. Design.* **6**:64–86.
15. **Gorelick, R. J., D. J. Chabot, D. E. Ott, T. D. Gagliardi, A. Rein, L. E. Henderson, and L. O. Arthur.** 1996. Genetic analysis of the zinc finger in the Moloney murine leukemia virus nucleocapsid domain: replacement of zinc-coordinating residues with other zinc-coordinating residues yields noninfectious particles containing genomic RNA. *J. Virol.* **70**:2593–2597.
16. **Gorelick, R. J., T. D. Gagliardi, W. J. Bosche, T. A. Wiltrout, L. V. Coren, D. J. Chabot, J. D. Lifson, L. E. Henderson, and L. O. Arthur.** 1999. Conservation of the retroviral nucleocapsid protein zinc finger is strongly influenced by its role in viral infection processes: characterization of HIV-1 particles containing mutant nucleocapsid zinc-coordinating sequences. *Virology* **256**:92–104.
17. **Graham, R. L., and M. R. Denison.** 2006. Replication of murine hepatitis virus is regulated by papain-like proteinase 1 processing of nonstructural proteins 1, 2, and 3. *J. Virol.* **80**:11610–11620.

18. **Graham, R. L., A. C. Sims, S. M. Brockway, R. S. Baric, and M. R. Denison.** 2005. The nsp2 replicase proteins of murine hepatitis virus and severe acute respiratory syndrome coronavirus are dispensable for viral replication. *J. Virol.* **79**:13399–13411.
19. **Gultyaev, A. P., F. H. van Batenburg, and C. W. Pleij.** 1995. The computer simulation of RNA folding pathways using a genetic algorithm. *J. Mol. Biol.* **250**:37–51.
20. **Guo, J., T. Wu, J. Anderson, B. F. Kane, D. G. Johnson, R. J. Gorelick, L. E. Henderson, and J. G. Levin.** 2000. Zinc finger structures in the human immunodeficiency virus type 1 nucleocapsid protein facilitate efficient minus- and plus-strand transfer. *J. Virol.* **74**:8980–8988.
21. **Harcourt, B. H., D. Jukneliene, A. Kanjanahaluethai, J. Bechill, K. M. Severson, C. M. Smith, P. A. Rota, and S. C. Baker.** 2004. Identification of severe acute respiratory syndrome coronavirus replicase products and characterization of papain-like protease activity. *J. Virol.* **78**:13600–13612.
22. **Kim, J. L., K. A. Morgenstern, C. Lin, T. Fox, M. D. Dwyer, J. A. Landro, S. P. Chambers, W. Markland, C. A. Lepre, E. T. O'Malley, S. L. Harbeson, C. M. Rice, M. A. Murcko, P. R. Caron, and J. A. Thomson.** 1996. Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell* **87**:343–355.
23. **Kummerer, B. M., and C. M. Rice.** 2002. Mutations in the yellow fever virus nonstructural protein NS2A selectively block production of infectious particles. *J. Virol.* **76**:4773–4784.
24. **Laity, J. H., B. M. Lee, and P. E. Wright.** 2001. Zinc finger proteins: new insights into structural and functional diversity. *Curr. Opin. Struct. Biol.* **11**:39–46.
25. **Landt, O., H.-P. Grunert, and U. Hahn.** 1990. A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene* **96**:125–128.
26. **Lee, N., R. J. Gorelick, and K. Musier-Forsyth.** 2003. Zinc finger-dependent HIV-1 nucleocapsid protein-TAR RNA interactions. *Nucleic Acids Res.* **31**:4847–4855.
27. **Liew, C. K., R. J. Y. Simpson, A. H. Y. Kwan, L. A. Crofts, F. E. Loughlin, J. M. Matthews, M. Crossley, and J. P. Mackay.** 2005. Zinc fingers as protein recognition motifs: structural basis for the GATA-1/Friend of GATA interaction. *Proc. Natl. Acad. Sci. USA* **102**:583–588.
28. **Lu, D., M. Alexandra Searles, and A. Klug.** 2003. Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. *Nature* **426**:96–100.
29. **Maclachlan, N. J., U. B. Balasuriya, J. F. Hedges, T. M. Schweidler, W. H. McCollum, P. J. Timoney, P. J. Hullinger, and J. F. Patton.** 1998. Serologic response of horses to the structural proteins of equine arteritis virus. *J. Vet. Diagn. Investig.* **10**:229–236.
30. **Matthews, J. M., K. Kowalski, C. K. Liew, B. K. Sharpe, A. H. Fox, M. Crossley, and J. P. Mackay.** 2000. A class of zinc fingers involved in protein-protein interactions: biophysical characterization of CCHC fingers from Fog and U-shaped. *Eur. J. Biochem.* **267**:1030–1038.
31. **McCarty, A. S., G. Kleiger, D. Eisenberg, and S. T. Smale.** 2003. Selective dimerization of a C2H2 zinc finger subfamily. *Mol. Cell* **11**:459–470.
32. **Mesters, J. R., J. Tan, and R. Hilgenfeld.** 2006. Viral enzymes. *Curr. Opin. Struct. Biol.* **16**:776–786.
33. **Miller, J., A. D. McLachlan, and A. Klug.** 1985. Repetitive zinc-binding domains in the protein transcription factor Iiia from *Xenopus* oocytes. *EMBO J.* **4**:1609–1614.
34. **Miller, W. A., and G. Koev.** 2000. Synthesis of subgenomic RNAs by positive-strand RNA viruses. *Virology* **273**:1–8.
35. **Molenkamp, R., B. C. D. Rozier, S. Greve, W. J. M. Spaan, and E. J. Snijder.** 2000. Isolation and characterization of an arterivirus defective interfering RNA genome. *J. Virol.* **74**:3156–3165.
36. **Molenkamp, R., H. van Tol, B. C. D. Rozier, Y. van der Meer, W. J. M. Spaan, and E. J. Snijder.** 2000. The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription. *J. Gen. Virol.* **81**:2491–2496.



37. **Newman, J. F., P. G. Piatti, B. M. Gorman, T. G. Burrage, M. D. Ryan, M. Flint, and F. Brown.** 1994. Foot-and-mouth disease virus particles contain replicase protein 3D. *Proc. Natl. Acad. Sci. USA* **91**:733–737.
38. **Pasternak, A. O., W. J. M. Spaan, and E. J. Snijder.** 2004. Regulation of relative abundance of arterivirus subgenomic rRNAs. *J. Virol.* **78**:8102–8113.
39. **Pasternak, A. O., W. J. M. Spaan, and E. J. Snijder.** 2006. Nidovirus transcription: how to make sense...? *J. Gen. Virol.* **87**:1403–1421.
40. **Pasternak, A. O., E. van den Born, W. J. M. Spaan, and E. J. Snijder.** 2001. Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. *EMBO J.* **20**:7220–7228.
41. **Pasternak, A. O., E. van den Born, W. J. M. Spaan, and E. J. Snijder.** 2003. The stability of the duplex between sense and antisense transcription-regulating sequences is a crucial factor in arterivirus subgenomic mRNA synthesis. *J. Virol.* **77**:1175–1183.
42. **Paul, A. V., J. H. van Boom, D. Filippov, and E. Wimmer.** 1998. Protein-primed RNA synthesis by purified poliovirus RNA polymerase. *Nature* **393**: 280–284.
43. **Pedersen, K. W., Y. van der Meer, N. Roos, and E. J. Snijder.** 1999. 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.* **73**:2016–2026.
44. **Peng, C.-W., A. J. Napuli, and V. V. Dolja.** 2003. Leader proteinase of *Beet yellows virus* functions in long-distance transport. *J. Virol.* **77**:2843–2849.
45. **Posthuma, C. C., D. D. Nedialkova, J. C. Zevenhoven-Dobbe, J. H. Blokhuis, A. E. Gorbalenya, and E. J. Snijder.** 2006. Site-directed mutagenesis of the nidovirus replicative endoribonuclease NendoU exerts pleiotropic effects on the arterivirus life cycle. *J. Virol.* **80**:1653–1661.
46. **Puustinen, P., M. L. Rajamaki, K. I. Ivanov, J. P. T. Valkonen, and K. Makinen.** 2002. Detection of the potyviral genome-linked protein VPg in virions and its phosphorylation by host kinases. *J. Virol.* **76**:12703–12711.
47. **Ramboarina, S., N. Moreller, M. C. Fournie-Zaluski, and B. P. Roques.** 1999. Structural investigation on the requirement of CCHH zinc finger type in nucleocapsid protein of human immunodeficiency virus 1. *Biochemistry* **38**:9600–9607.
48. **Sawicki, S. G., and D. L. Sawicki.** 1995. Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. *Adv. Exp. Biol. Med.* **380**:499–506.
49. **Sawicki, S. G., and D. L. Sawicki.** 2005. Coronavirus transcription: a perspective, p. 31–55. *In* L. Enjuanes (ed.), *Coronavirus replication and reverse genetics*. Springer, Berlin, Germany.
50. **Sawicki, S. G., D. L. Sawicki, and S. G. Siddell.** 2007. A contemporary view of coronavirus transcription. *J. Virol.* **81**:20–29.
51. **Seybert, A., C. C. Posthuma, L. C. van Dinten, E. J. Snijder, A. E. Gorbalenya, and J. Ziebuhr.** 2005. A complex zinc finger controls the enzymatic activities of nidovirus helicases. *J. Virol.* **79**:696–704.
52. **Siddell, S. G., J. Ziebuhr, and E. J. Snijder.** 2005. Coronaviruses, toroviruses, and arteriviruses, p. 823–856. *In* B. W. Mahy and V. ter Meulen, Topley and Wilson's microbiology and microbial infections: virology volume. Hodder Arnold, London, United Kingdom.
53. **Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. M. Poon, Y. Guan, M. Rozanov, W. J. M. Spaan, and A. E. Gorbalenya.** 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J. Mol. Biol.* **331**:991–1004.
54. **Snijder, E. J., and J. J. M. Meulenberg.** 1998. The molecular biology of arteriviruses. *J. Gen. Virol.* **79**:961–979.

55. **Snijder, E. J., and J. J. M. Meulenberg.** 2001. Arteriviruses, p. 1205–1220. *In* D. M. Knipe and P. M. Howley, *Fields virology*. Lippincott, Williams & Wilkins, Philadelphia, PA.
56. **Snijder, E. J., S. G. Siddell, and A. E. Gorbalenya.** 2005. The order Nidovirales, p. 390–404. *In* B. W. Mahy and V. ter Meulen (ed.), *Topley and Wilson's microbiology and microbial infections: virology volume*. Hodder Arnold, London, United Kingdom.
57. **Snijder, E. J., H. van Tol, N. Roos, and K. W. Pedersen.** 2001. Non-structural proteins 2 and 3 interact to modify host cell membranes during the formation of the arterivirus replication complex. *J. Gen. Virol.* **82**:985–994.
58. **Snijder, E. J., A. L. M. Wassenaar, and W. J. M. Spaan.** 1992. The 5' end of the equine arteritis virus replicase gene encodes a papainlike cysteine protease. *J. Virol.* **66**:7040–7048.
59. **Snijder, E. J., A. L. M. Wassenaar, and W. J. M. Spaan.** 1994. Proteolytic processing of the replicase ORF1a protein of equine arteritis virus. *J. Virol.* **68**:5755–5764.
60. **Snijder, E. J., A. L. M. Wassenaar, W. J. M. Spaan, and A. E. Gorbalenya.** 1995. The arterivirus nsp2 protease: an unusual cysteine protease with primary structure similarities to both papain-like and chymotrypsin-like proteases. *J. Biol. Chem.* **270**:16671–16676.
61. **Stote, R. H., E. Kellenberger, H. Muller, E. Bombarda, B. P. Roques, B. Kieffer, and Y. Mely.** 2004. Structure of the His44Ala single point mutant of the distal finger motif of HIV-1 nucleocapsid protein: a combined NMR, molecular dynamics simulation, and fluorescence study. *Biochemistry* **43**: 7687–7697.
62. **Tijms, M. A., L. C. van Dinten, A. E. Gorbalenya, and E. J. Snijder.** 2001. A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus. *Proc. Natl. Acad. Sci. USA* **98**:1889–1894.
63. **van Aken, D., J. C. Zevenhoven-Dobbe, A. E. Gorbalenya, and E. J. Snijder.** 2006. Proteolytic maturation of replicase polyprotein pp1a by the nsp4 main proteinase is essential for equine arteritis virus replication and includes internal cleavage of nsp7. *J. Gen. Virol.* **87**:3473–3482.
64. **van den Born, E., A. P. Gultyaev, and E. J. Snijder.** 2004. Secondary structure and function of the 5'-proximal region of the equine arteritis virus RNA genome. *RNA* **10**:424–437.
65. **van den Born, E., C. C. Posthuma, A. P. Gultyaev, and E. J. Snijder.** 2005. Discontinuous subgenomic RNA synthesis in arteriviruses is guided by an RNA hairpin structure located in the genomic leader region. *J. Virol.* **79**: 6312–6324.
66. **van den Born, E., C. C. Posthuma, K. Knoops, and E. J. Snijder.** 2007. An infectious recombinant equine arteritis virus expressing green fluorescent protein from its replicase gene. *J. Gen. Virol.* **88**:1196–1205.
67. **van der Meer, Y., H. van Tol, J. Krijnse Locker, and E. J. Snijder.** 1998. ORF1a-encoded replicase subunits are involved in the membrane association of the arterivirus replication complex. *J. Virol.* **72**:6689–6698.
68. **van Dinten, L. C., J. A. den Boon, A. L. M. Wassenaar, W. J. M. Spaan, and E. J. Snijder.** 1997. An infectious arterivirus cDNA clone: identification of a replicase point mutation which abolishes discontinuous mRNA transcription. *Proc. Natl. Acad. Sci. USA* **94**:991–996.
69. **van Dinten, L. C., H. van Tol, A. E. Gorbalenya, and E. J. Snijder.** 2000. The predicted metal-binding region of the arterivirus helicase protein is involved in subgenomic mRNA synthesis, genome replication, and virion biogenesis. *J. Virol.* **74**:5213–5223.
70. **van Marle, G., J. C. Dobbe, A. P. Gultyaev, W. Luytjes, W. J. M. Spaan, and E. J. Snijder.** 1999. Arterivirus discontinuous mRNA transcription is guided by base-pairing between sense and antisense transcription-regulating sequences. *Proc. Natl. Acad. Sci. USA* **96**:12056–12061.

71. **Wieringa, R., A. A. F. de Vries, J. van der Meulen, G.-J. Godeke, J. J. M. Onderwater, H. van Tol, H. K. Koerten, A. M. Mommaas, E. J. Snijder, and P. J. M. Rottier.** 2004. Structural protein requirements in equine arteritis virus assembly. *J. Virol.* **78**:13019–13027.
72. **Zevenhoven-Dobbe, J. C., S. Greve, H. van Tol, W. J. M. Spaan, and E. J. Snijder.** 2004. Rescue of disabled infectious single-cycle (DISC) equine arteritis virus by using complementing cell lines that express minor structural glycoproteins. *J. Gen. Virol.* **85**:3709–3714.
73. **Ziebuhr, J.** 2004. Molecular biology of severe acute respiratory syndrome coronavirus. *Curr. Opin. Microbiol.* **7**:412–419.
74. **Ziebuhr, J., B. Schelle, N. Karl, E. Minskaia, S. Bayer, S. G. Siddell, A. E. Gorbalenya, and V. Thiel.** 2007. Human coronavirus 229E papain-like proteases have overlapping specificities but distinct functions in viral replication. *J. Virol.* **81**:3922–3932.
75. **Ziebuhr, J., E. J. Snijder, and A. E. Gorbalenya.** 2000. Virus-encoded proteinases and proteolytic processing in the *Nidovirales*. *J. Gen. Virol.* **81**:853–879.
76. **Ziebuhr, J., V. Thiel, and A. E. Gorbalenya.** 2001. The autocatalytic release of a putative RNA virus transcription factor from its polyprotein precursor involves two paralogous papain-like proteases that cleave the same peptide bond. *J. Biol. Chem.* **276**:33220–33232.
77. **Zuker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. *Science* **244**:48–52.
78. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**:3406–3415.



# Chapter 6

**Arterivirus nsp1 modulates the  
accumulation of minus-strand  
templates to control the relative  
abundance of viral mRNAs**

Danny D. Nedialkova,  
Alexander E. Gorbalenya, and  
Eric J. Snijder

PLoS Pathog. 2010. 6(2): e1000772.

## ABSTRACT

The gene expression of plus-strand RNA viruses with a polycistronic genome depends on translation and replication of the genomic mRNA, as well as synthesis of subgenomic (sg) mRNAs. Arteriviruses and coronaviruses, distantly related members of the nidovirus order, employ a unique mechanism of discontinuous minus-strand RNA synthesis to generate subgenome-length templates for the synthesis of a nested set of sg mRNAs. Non-structural protein 1 (nsp1) of the arterivirus equine arteritis virus (EAV), a multifunctional regulator of viral RNA synthesis and virion biogenesis, was previously implicated in controlling the balance between genome replication and sg mRNA synthesis. Here, we employed reverse and forward genetics to gain insight into the multiple regulatory roles of nsp1. Our analysis revealed that the relative abundance of viral mRNAs is tightly controlled by an intricate network of interactions involving all nsp1 subdomains. Distinct nsp1 mutations affected the quantitative balance among viral mRNA species, and our data implicate nsp1 in controlling the accumulation of full-length and subgenome-length minus-strand templates for viral mRNA synthesis. The moderate differential changes in viral mRNA abundance of nsp1 mutants resulted in similarly altered viral protein levels, but progeny virus yields were greatly reduced. Pseudorevertant analysis provided compelling genetic evidence that balanced EAV mRNA accumulation is critical for efficient virus production. This first report on protein-mediated, mRNA-specific control of nidovirus RNA synthesis reveals the existence of an integral control mechanism to fine-tune replication, sg mRNA synthesis, and virus production, and establishes a major role for nsp1 in coordinating the arterivirus replicative cycle.

abstract

## INTRODUCTION

Plus-strand RNA (+RNA) viruses are ubiquitous pathogens of plants, animals, and humans. The translation of their messenger-sense RNA genome yields the core viral enzymes that always include an RNA-dependent RNA polymerase (RdRp) and assemble into a cytoplasmic machinery for viral RNA synthesis. Many +RNA virus groups employ polycistronic genomes and different mechanisms to express genes located downstream of the 5'-proximal open reading frame (ORF). One of these mechanisms involves the synthesis of subgenomic (sg) mRNAs (referred to as "transcription" in this paper). Although the sg mRNAs of +RNA viruses are invariably 3'-coterminally with the viral genome, diverse +RNA viruses have evolved different mechanisms for their production<sup>18</sup>.

The order *Nidovirales* comprises several clades of distantly related enveloped +RNA viruses, including the arteri- and coronavirus families, which infect a wide variety of hosts, ranging from invertebrates to humans. Human coronaviruses are associated with respiratory disease (including severe acute respiratory syndrome (SARS), reviewed in reference<sup>27</sup>) and arteriviruses like porcine reproductive and respiratory syndrome virus (PRRSV) are important veterinary pathogens. Members of the nidovirus order are characterized by their exceptional genetic complexity, and the group includes the virus families with the largest RNA genomes described to date (25-32 kb). Nidoviruses share important traits in their genome organization and gene expression mechanisms, and their key replicative enzymes are presumed to be evolutionarily related (for a review, see reference<sup>13</sup>). Their polycistronic genomes are 5'-capped, 3'-polyadenylated, and the two 5'-most open reading frames (ORFs) – ORF1a and ORF1b, encode the viral replicase subunits segregated in two large replicase polyproteins, pp1a and pp1ab, the expression of the latter controlled by a -1 ribosomal frameshift (Fig. 1A). Autoproteolytic processing of these precursors generates between 13 and 16 non-structural proteins (nsps) that direct viral RNA synthesis. Besides genome replication, arteri- and coronavirus RdRp-containing complexes also mediate the synthesis of a distinctive nested set of sg mRNAs that are both 5'- and 3'-coterminally with the viral genome and hence consist of sequences that are noncontiguous in the genomic RNA (Fig. 1B).

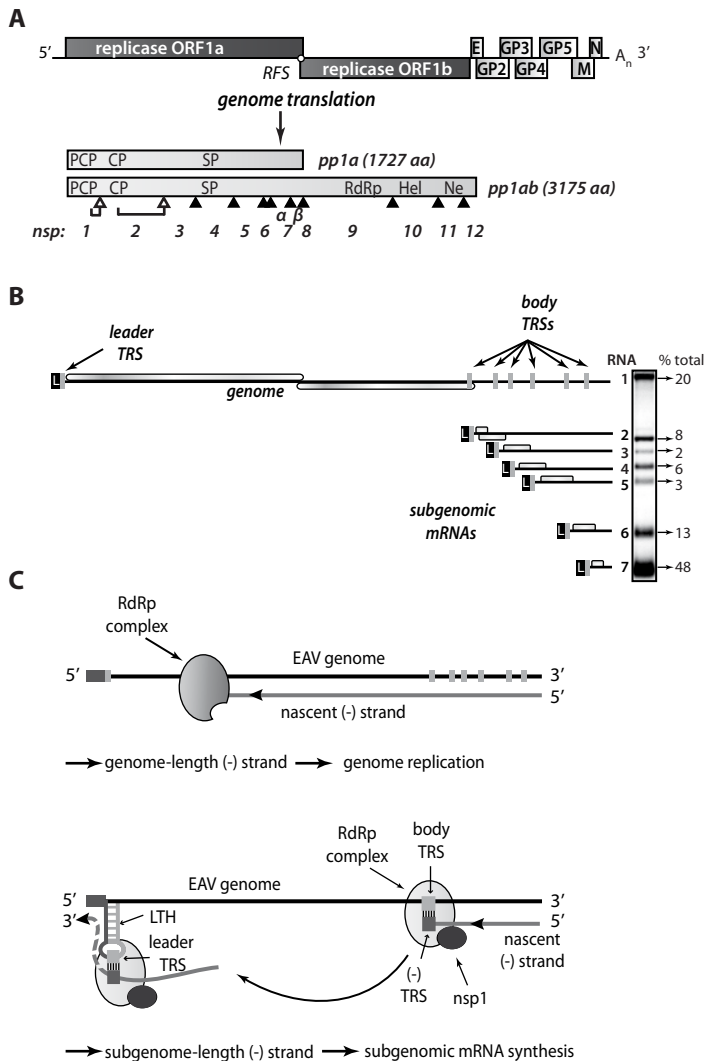
Despite recent advances in the structural and functional characterization of individual replicase subunits, the molecular details of nidovirus replication and gene expression remain poorly understood. Studies with nidovirus model systems such as equine arteritis virus (EAV), the arterivirus prototype, have provided some critical insights about viral replicase functions and the regulation of RNA synthesis in infected cells. EAV replicase pp1a and pp1ab are co- and post-translationally cleaved into 13 nsps by viral proteases residing in nsp1, nsp2, and nsp4. The seven viral structural proteins, which are all dispensable for replication and transcription<sup>20</sup>, are encoded in a set of overlapping ORFs located in the 3'-proximal quarter of the genome (Fig. 1A). In the six sg mRNAs used to express these ORFs, a common "leader" sequence representing the 5'-proximal 206 nucleotides of the genome is linked to different "body" segments that are co-linear with the 3'-proximal part of the genome (Fig. 1B).



According to the widely supported model proposed by Sawicki and Sawicki<sup>30</sup> (Fig. 1C), the structure of the arterivirus and coronavirus sg mRNAs derives from a discontinuous step during minus-strand RNA synthesis, which is guided by specific RNA signals and resembles copy-choice RNA recombination<sup>24,47,55</sup>. Conserved transcription-regulating sequences (TRS; core sequence 5'UCAACU 3' in EAV) precede each structural protein ORF (body TRSs). The same sequence motif is also present at the 3'-end of the genomic leader sequence (leader TRS). Minus-strand RNA synthesis, initiated at the 3'-end of the viral genome, is presumably attenuated at one of the body TRS regions (Fig. 1C; reviewed in references<sup>22,31</sup>). Subsequently, the nascent minus strand, carrying the body TRS complement at its 3' end, is translocated to the 5'-proximal region of the genomic template. During this step, the genomic leader TRS serves as a base-pairing target for the 3' end of the nascent minus strand, a role that is facilitated, in the case of EAV, by its presence in the loop of an RNA hairpin<sup>42</sup>. When minus strand synthesis resumes, nascent strands are extended with the complement of the genomic leader sequence, yielding a nested set of subgenome-length minus-strand templates that are used for the subsequent synthesis of the various sg mRNAs. If attenuation does not occur, minus-strand RNA synthesis proceeds to yield a full-length complement of the genome, the intermediate required for its replication.

Clearly, the protein and RNA factors that determine whether a nidovirus RdRp complex operates in continuous or discontinuous mode, i.e. produces a full-length or a subgenome-length minus strand, must be critical for the coordination of the nidovirus replicative cycle. As in other nidoviruses, the EAV genomic RNA (RNA1) and sg mRNAs (RNA2 – RNA7) accumulate in specific molar ratios (see Fig. 1B) that are essentially constant until the peak of viral RNA synthesis is reached<sup>11</sup>. The relative abundance of the transcripts is presumably dictated by the "attenuation rate" at each of the successive body TRSs encountered during minus-strand synthesis, which is primarily determined by the base-pairing potential between the leader TRS and the body TRS complement in the nascent minus strand. Also the sequence context of body TRS motifs and their proximity to the genomic 3' end, which is reflected in the "gradient" of sg RNA sizes, can influence the accumulation of viral RNA species. The importance of TRS-driven RNA-RNA interactions and the potential for a regulatory role of higher order RNA structures was outlined above (reviewed in references<sup>22,31</sup>). At the protein level, however, only a single nidovirus protein specifically involved in transcription was identified thus far: EAV nsp1 was found to be essential for sg mRNA production, while being dispensable for genome replication<sup>37,39</sup>. Remarkably, nsp1 is also the first protein expressed during infection: it is co-translationally released from the nascent replicase polyproteins by a papain-like cysteine proteinase activity (PCP $\beta$ ) in its C-terminal domain (Fig 2). Comparative sequence analysis identified two additional conserved domains: a second, proteolytically silent PCP domain that is functional in other arteriviruses<sup>9</sup> (PCP $\alpha$ ), and an N-terminal zinc finger (ZF) domain<sup>36,39</sup> that is critical for transcription and efficient production of infectious progeny<sup>37,39</sup>. Since the accumulation of all sg mRNAs was blocked in the absence of nsp1, the protein was proposed to control a switch between replication and transcription<sup>39</sup>.





**Figure 1.** Organization and expression of the polycistronic EAV +RNA genome. (A) Top: EAV genome organization, showing the 5'-proximal replicase open reading frames (ORFs), as well as the downstream ORFs encoding the viral structural proteins envelope (E), membrane (M), nucleocapsid (N), and glycoproteins (GP) 2-5 and the 3' poly(A) tail ( $A_n$ ). Bottom: overview of the pp1a and pp1ab replicase polyproteins that result from genome translation, which requires an ORF1a/1b ribosomal frameshift (RFS) to produce pp1ab. Arrowheads represent sites cleaved by the three virus-encoded proteases (open for autoproteolytically processed ones, closed for sites processed by the main proteinase in nsp4). The resulting nonstructural proteins (nsp) are numbered. The key viral enzymatic domains such as the nsp1 papain-like cysteine proteinase  $\beta$  (PCP), nsp2 cysteine proteinase (CP), nsp4 serine proteinase (SP), nsp9 viral RNA-dependent RNA polymerase (RdRp), nsp10 helicase (Hel), and nsp11 endoribonuclease (Ne) are indicated. (B) Overview of viral mRNA species produced in EAV-infected cells. The ORFs expressed from the respective mRNAs are shown in, and the 5' leader sequence (L) is depicted. The positions of transcription-regulating sequences (TRS) are indicated. The gel hybridization image on the right is representative of the wild-type accumulation levels of the seven EAV mRNAs at the time point used for analysis in the study (see text for details). The amount of each mRNA, determined by quantitative phosphorimager analysis, is

indicated as percentage of the total amount of viral mRNA. (C) Model for EAV replication and transcription. Continuous minus-strand RNA synthesis yields a genome-length minus strand template for genome replication, a process for which nsp1 is dispensable. Discontinuous minus-strand RNA synthesis results in a nested set of subgenome-length minus strands that serve as templates for sg mRNA synthesis (see text for details). Nsp1 is crucial for this process, which is also guided by a base pairing interaction between the TRS complement [(-)TRS] at the 3' end of the nascent minus-strand and the genomic leader TRS, present in a RNA hairpin structure (LTH).

We have now explored the key regulatory roles of nsp1 in the EAV replicative cycle in unprecedented detail. Our results indicate that in addition to the ZF region, both PCP subdomains of nsp1 are essential for transcription, and suggest an additional role of PCP $\alpha$  in virus production. We also established that nsp1 modulates viral RNA accumulation in an mRNA-specific manner, and thus maintains the balance among the seven viral mRNAs, including the genome. Our data suggests that nsp1 does so by controlling the levels of the full-length and subgenome-length minus-strand templates required for viral mRNA synthesis. The results we obtained from detailed characterization of nsp1 mutants and pseudorevertants provided compelling evidence for a close link between the regulation of individual nidovirus mRNA levels and the efficient production of infectious progeny.

## **MATERIALS AND METHODS**

### **Cell lines**

Baby hamster kidney cells (BHK-21; ATCC CCL10) were used for all experiments. The cells were maintained at 37°C in BHK-21 medium (Glasgow MEM; Invitrogen) supplemented with 5% fetal calf serum (FCS), 10% tryptose phosphate broth, 100 U/ml of penicillin, 100  $\mu$ g/ml of streptomycin and 10 mM HEPES, pH=7.4. Upon transfection or infection with wt or mutant EAV, BHK-21 cells were incubated at 39.5°C, since this elevated temperature shortens the replication time of the virus substantially without any adverse side effects<sup>44</sup>.

### **EAV reverse genetics**

The substitutions in the nsp1-coding sequence listed in Table 1 and Table 3 were engineered using appropriate shuttle vectors and standard site-directed mutagenesis PCR<sup>16</sup>. Sequence analysis of the cloned fragments was used to verify the introduction of the appropriate nucleotide substitutions and exclude the presence of undesired mutations. The mutations in the nsp1-coding sequence were then transferred to pEAV211 or pEAN551, both derivatives of EAV full-length cDNA clone pEAV030<sup>44</sup> containing some engineered restriction sites. Viruses derived from either pEAV211 or pEAN551 were previously shown to display a wt phenotype<sup>28,41</sup>. The virus derived from the pEAV211 construct was used as a wt control in all experiments.

In vitro RNA transcription from *Xho*I-linearized wt or mutant EAV full-length cDNA clones was performed using the mMESSAGE mMACHINE T7 Kit (Ambion). Seven  $\mu$ g of *in vitro*-synthesized EAV RNA were electroporated into  $3.5 \times 10^6$  BHK-21 cells using the Amaxa Cell Line Nucleofector Kit T and the program T-020 of the Amaxa Nucleofector (Lonza) according to the manufacturer's instructions. Cells were resuspended in BHK-21 medium and subsequently seeded on coverslips for immunofluorescence analysis or in 6-well clusters for analysis of intracellular protein and RNA levels, as well as virus production.

### **EAV infection and plaque assays**

For EAV infection, subconfluent monolayers of BHK-21 cells were inoculated with transfected cell culture supernatant diluted in PBS-DEAE/2% FCS. Following incubation at 39.5°C for 1 h, the inoculum was removed, DMEM/2% FCS was added, and the cells were incubated at 39.5°C for 16-18 h. For virus titration, BHK-21 cells seeded in 6-well clusters were infected with serial 10-fold dilutions of supernatants harvested from transfected cells and then incubated under semi-solid overlays consisting of DMEM supplemented with 50 mM HEPES, pH=7.4, 2% FCS and 1.2% Avicel (FMC BioPolymer) at 39.5°C for 72 h. The overlays were aspirated, cells were fixed with 8% formaldehyde in PBS, and stained with crystal violet. For plaque purification, a solid overlay of DMEM containing 50 mM HEPES, pH=7.4, 2% FCS and 1% agarose was used.

### **Immunofluorescence analysis**

Immunofluorescence analysis of EAV-transfected cells was performed as described previously<sup>43</sup>. Briefly, cells were analyzed at 11 hour post-transfection by dual labeling with a rabbit antiserum recognizing EAV nsp3<sup>25</sup> and an anti-N mouse monoclonal antibody<sup>17</sup>. These proteins are expressed from RNA1 and RNA7, respectively. Nuclei were visualized for cell counting by staining with 1  $\mu$ g/ml Hoechst 33258 (Sigma-Aldrich). Transfection efficiencies were determined at 11 h post-transfection by counting cells with the Scion Image software (Scion Corporation) and calculating the percentage of cells positive for EAV nsp3.

### **RNA isolation and detection of EAV mRNAs by gel hybridization**

Analysis of viral RNA accumulation was carried out before completion of the first replication cycle. Cells transfected with wt or mutant EAV derivatives were lysed at 11 h post-transfection, and total intracellular RNA was isolated by acid phenol extraction as previously described<sup>48</sup>. Viral mRNAs were detected by resolving total RNA in denaturing agarose-formaldehyde gels, and equal sample loading was confirmed by ethidium bromide staining of ribosomal RNA. The gels were subsequently dried and hybridized to a <sup>32</sup>P-labeled probe (E154, 5'-TTGGTTCCTGGGTGGCTAATAACTACTT-3') complementary to

the 3' end of the genome that recognizes both genome and all sg mRNAs<sup>48</sup>. The gels were exposed to phosphorimager screens, which were subsequently scanned using a Typhoon Variable Mode Imager (GE Healthcare). Image analysis and quantification of band intensities were performed with the ImageQuant TL software (GE Healthcare).

### **Detection of EAV minus-strand RNAs by ribonuclease protection assays**

A two-cycle ribonuclease (RNase) protection assay adapted from references<sup>21,29</sup> was used for the detection of EAV minus-strand RNAs. Total intracellular RNA isolated at 11 h post-transfection was dissolved in 10  $\mu$ l of Hybridization Buffer III (RPA III Kit; Ambion), denatured for 3 min at 95°C and incubated for 16 h at 55°C. Samples were then digested with 5 U of RNase T1 (Ambion) per  $\mu$ g total RNA in 10 mM Tris pH=7.5, 300 mM NaCl, 5 mM EDTA for 60 min at ambient temperature. Following proteinase K treatment and phenol:chloroform extraction, 0.5  $\mu$ g of yeast RNA per  $\mu$ g total RNA was added as a carrier. After ethanol precipitation, equal amounts of a radiolabelled probe (see below) were added to each sample and, following sample denaturation at 85°C for 5 min, hybridization was carried out at 55°C for 16 h. RNase digestion of unhybridized RNA was performed using the RPA III Kit according to the manufacturer's protocol. Protected fragments were resolved in 5% polyacrylamide/8M urea gels, which were dried and exposed to phosphorimager screens. Image analysis and quantification were performed as described above.

To generate probes for minus-strand detection, cDNA fragments derived from the EAV genome (nucleotides (nt) 3687 – 4013), RNA6 (nt 68-206 from the leader sequence and nt 11,870-12,057 from the body sequence) and RNA7 (nt 68-206 and nt 12,252-12,429) were inserted downstream of the T7 promoter in pcDNA3.1 using standard cloning procedures. Radiolabelled RNA transcripts were generated by *in vitro* transcription in the presence of [ $\alpha$ -<sup>32</sup>P]CTP (Perkin Elmer) using MAXIscript T7 Kit (Ambion) and purified from 5% polyacrylamide/8M urea gels by elution for 3h at 37°C in 0.5 M NH<sub>4</sub>OAc, 0.2% SDS, 1mM EDTA. The transcript generated for detection of genomic minus strands, pRNA1, was 356 nt long and contained 327 nt of (-)RNA1-specific sequence. The transcript and EAV-specific sequence length were 382 nt and 327 nt, respectively, for the probe detecting (-)RNA6 (pRNA6), and 372 nt and 319 nt for the probe specific for (-)RNA7 (pRNA7). Detection of RNA1(-) was performed using sample RNA corresponding to approximately 1.25x10<sup>4</sup> cells and 20 fmol of radiolabelled probe. Levels of (-)RNA6 and (-)RNA7 were determined in samples corresponding to 4x10<sup>4</sup> and 2.5 x10<sup>4</sup> cells, respectively, using 5 fmol radiolabelled probe. These conditions ensured that the values obtained were in the linear range of the assays (data not shown).

### **Protein analysis**

Cells transfected with wt or mutant EAV derivatives were lysed at 11 h post-transfection as described previously<sup>7</sup>. The protein concentration in the lysates was determined using the Bio-Rad protein assay reagent. Equal amounts of total protein were subjected to SDS-PAGE

and transferred to Hybond-P PVDF membrane (GE Healthcare) by semidry blotting. After blocking with 5% non-fat milk in PBS containing 0.5% Tween-20, the membranes were incubated with the following antibodies: anti-EAV nsp3 (see above), rabbit anti-M<sup>7</sup>, anti-N (see above), or an anti- $\beta$ -actin mouse monoclonal antibody (AC-74, Sigma), all diluted in PBS containing 5% non-fat milk, 0.5% bovine serum albumin and 0.5% Tween-20. HRP-conjugated secondary antibodies (DAKO) and an ECL-Plus kit (GE Healthcare) were used for detection.

## Reverse transcription and quantitative PCR

In order to determine relative specific infectivity, supernatants from BHK-21 cells transfected with wt or mutant EAV derivatives were harvested 24 h after transfection and clarified by low-speed centrifugation. Virions from 1 ml of clarified supernatant were purified by pelleting through a 0.4 ml cushion of 20 % sucrose in 20 mM Tris pH=7.5, 100 mM NaCl, 1 mM EDTA at 55,000 rpm for 45 min at 4°C using a TLS-55 rotor in a Beckman tabletop ultracentrifuge. Virion RNA was isolated from pellet fractions by acid phenol extraction as described above. Complementary DNAs were synthesized with Thermoscript Reverse Transcriptase (Invitrogen) using a primer complementary to a region in ORF1a of the EAV genome (EAV418as, 5'-AGCCGCACCTTCACATTG-3'). Quantitative PCR (qPCR) was performed essentially as previously described<sup>132,49</sup>. Briefly, a cDNA aliquot was amplified with EAV-specific oligonucleotides EAV418as and EAV417s (5' CATCTCTTGCTTGCTCCTTAG-3') using HotStar *Taq* Polymerase (Qiagen) and SYBR Green I (Molecular Probes) in an iCycler machine (Bio-Rad). The data obtained were analyzed with iCycler software, and the specificity of the reaction was confirmed by the melting curve of the amplified products. To generate a standard curve, serial ten-fold dilutions of the virion RNA sample derived from cells transfected with the wt EAV construct were reverse-transcribed and amplified by qPCR in parallel. The resulting standard curve had an  $R^2=0.99$  and a 6-log linear range for the EAV ORF1a amplicon (data not shown). The relative genomic RNA contents of virions produced by EAV mutants were calculated by comparing their threshold cycle (Ct) values against the standard curve and the resulting values were normalized to the wt genomic RNA content, which was set at 1. The relative specific infectivity of each EAV mutant was then determined by dividing the respective mutant:wt pfu ratio by the mutant:wt relative genomic RNA content.

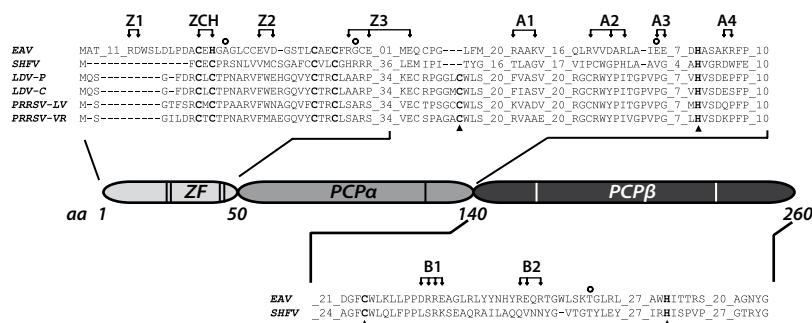
## RESULTS

### Rationale for mutagenesis of EAV nsp1

Previous studies of the role of nsp1 in the EAV replicative cycle focused on the conserved amino acids presumed to be essential either for zinc binding by the ZF domain or for the catalytic activity of the PCP $\beta$  autoprotease<sup>37</sup>. Mutations that blocked the release of nsp1 from the replicase polyproteins were lethal, likely due to their interference with downstream polyprotein processing steps that are essential for genome replication<sup>28</sup>.

<sup>40, 45</sup>. By contrast, replacements of putative zinc-coordinating residues either selectively abolished transcription of all viral sg mRNAs or interfered with virus production without affecting viral mRNA accumulation. In an attempt to expand our repertoire of viable nsp1 mutants, we now used two approaches: i) alanine scanning mutagenesis of non-conserved clusters of polar residues found throughout the nsp1 sequence, and ii) a Cys↔His interchange at the positions of residues Cys-25 and His-27, which have both been implicated in zinc coordination<sup>36,39</sup>. The first approach is less likely to perturb the protein's overall stability, since clusters of charged residues are usually found on the protein surface, where they may mediate interactions with other biomolecules via electrostatic interactions or hydrogen bond formation<sup>3,6,51</sup>. We reasoned that the second approach might preserve zinc coordination but could nevertheless have a subtle effect on zinc binding that might be translated in a measurable effect on one or more of nsp1's functions. Moreover, if these substitutions would compromise virus replication, isolation of revertant viruses encoding compensatory second-site mutations might reveal potential regulatory protein-protein or protein-RNA interactions.

Table 1 lists the nsp1 mutants characterized in this study. In all three subdomains of nsp1, clusters of two or three charged amino acids within a five- to seven-amino acid stretch were substituted with Ala. Constructs with Ala replacements in the ZF domain were designated Z (1 to 3), while those with replacements in the PCPα and PCPβ domains were designated A (1 to 4) and B (1 and 2), respectively (see Table 1 and Fig. 2). In addition, we swapped the Cys-25 and His-27 residues to generate the ZCH mutant. Full-length RNA transcribed from EAV cDNA clones encoding these nsp1 mutations was transfected into BHK-21 cells. Analysis of nsp1 mutant phenotypes was performed during the peak of viral RNA synthesis and before the bulk of infectious progeny was produced by the wild-type



**Figure 2.** Domain organization of EAV nsp1. The partial sequence alignment shows key regions in the three subdomains previously identified in the arterivirus nsp1 region. GenBank accession numbers for the full-length arterivirus genomes used for the alignment are as follows: EAV, NC\_002532; simian hemorrhagic fever virus (SHFV), NC\_003092; lactate dehydrogenase-elevating virus (LDV-P and LDV-C), NC\_001639 and NC\_002534; PRRSV-LV, M96262.2; PRRSV-VR, AY150564. Zinc-coordinating residues are indicated in bold font; the active-site Cys and His of PCPα and PCPβ are indicated with triangles (note the loss of the active-site Cys in EAV PCPα). The positions of amino acid clusters mutated in this study are indicated with arrows. All substitutions were with Ala, with the exception of the ZCH construct, in which Cys-25 and His-27 were swapped. The positions of mutations found in pseudorevertants are indicated with open circles.

(wt) control (11 h post-transfection; referred to as first-cycle analysis). Previous studies of intracellular viral RNA levels had been hampered by the considerable variability in transfection efficiencies of synthetic EAV full-length RNAs. For comparison between mutants, these studies used the genomic RNA as an internal standard for each sample to calculate relative ratios of viral mRNA accumulation levels<sup>24,37</sup>. Using an improved electroporation protocol (for details, see Materials and Methods), we now achieved very consistent and relatively high RNA transfection efficiencies (between 45% and 55% of positive cells at 11 h post-transfection with replication-competent synthetic EAV RNAs; data not shown). This allowed for the comparison of the absolute levels of mRNA accumulation and the detailed first-cycle analysis of EAV nsp1 mutants at a time point at which differences in virus production or (pseudo)reversion would not influence the assessment of their phenotype.

### All nsp1 subdomains are critical for transcription, while the ZF and PCPα domains are also important for virus production

The ZF domain of nsp1 is essential for sg mRNA production<sup>37,39</sup>, but the question of whether the PCPα and PCPβ domains also contribute to the protein's function in transcription has not been previously addressed, partly due to the nonviable phenotype of PCPβ mutants in which the nsp1/2 cleavage was impaired<sup>37</sup>. Consequently, we first analyzed the impact of clustered charged-to-alanine replacements in nsp1 on viral mRNA accumulation. Cells transfected with nsp1 mutants were harvested 11 h after transfection.

**Table 1.** Overview of the genotype and first-cycle phenotype of EAV nsp1 mutants described in this study

Construct	Genotype	Mutant codons <sup>a</sup>	Replication <sup>b</sup>	Transcription <sup>b</sup>
pEAV211	wt	NA <sup>c</sup>	+++	+++
Z1	R15A D16A	267- <b>GCC</b> GCC-272	+++++	-
ZCH	C25H H27C	297- <b>CAU</b> UGU-305	+++	++++
Z2	E34A D36A	324- <b>GCA</b> GCC-332	+++	+++
Z3	R46A E49A E52A	360- <b>GCC</b> GCA GCG-380	+++++	-
A1	R80A K83A	462- <b>GCA</b> GCA -474	+++++	++
A2	R103A D106A R108A	531- <b>GCU</b> GCA GCG-548	+++++	-
A3	E112A E113A	558- <b>GCC</b> GCG-563	+++	+++
A4	K126A R127A	600- <b>GCA</b> GCU-605	++++	+++
B1	D172A R173A R174A E175A	738- <b>GCU</b> GCA GCU GCG-749	-	-
B2	R186A E187 R189A	780- <b>GCC</b> GCA GCG-791	+++++	-

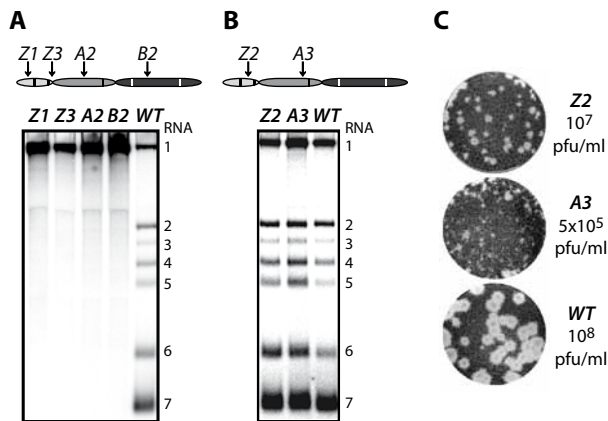
<sup>a</sup> Nucleotide substitutions are indicated in bold italics; numbers indicate the start and end coordinates of the mutated codons in the EAV genome.

<sup>b</sup> Transfected cells were analyzed by IFA and gel hybridization analysis at 11 h post-transfection. Wild-type levels of replication and transcription (based on accumulation levels of genomic and sg mRNAs, respectively) are indicated with +++. A two- to fourfold increase in genomic RNA levels as compared to wild-type is denoted by +++++, while an increase of more than fourfold is shown as ++++++. Likewise, a more than twofold increase or decrease in accumulation levels of at least two sg mRNA species is shown as +++++ and ++, respectively (see text for details)

<sup>c</sup> NA, not applicable.

tion, intracellular RNA was isolated and resolved in denaturing gels, and viral mRNAs were detected by hybridization to a probe complementary to the 3' end of the genome and thus recognizing all viral mRNAs. Substitutions in the ZF domain (mutant Z1), as well as in the region connecting the ZF and PCP $\alpha$  domains (Z3), the PCP $\alpha$  domain itself (A2) and, notably, also the PCP $\beta$  domain (B2) rendered viral sg mRNAs undetectable. In addition, all four mutants displayed a noticeable increase in genomic RNA levels (Fig. 3A). By contrast, accumulation of all viral mRNAs was blocked in the B1 mutant (Table 1 and data not shown), possibly due to the proximity of the charged cluster to the active site Cys of PCP $\beta$  (Fig. 2). These results demonstrate that all subdomains of nsp1, including PCP $\beta$ , are important for transcriptional control. This novel role of the PCP $\beta$  domain seems to be genetically separable from its autoproteolytic activity.

We previously reported that certain substitutions of proposed nsp1 zinc-coordinating residues considerably reduced the yield of infectious progeny without noticeably affecting viral RNA accumulation<sup>37</sup>. This phenotype was also observed in this study for mutants Z2 (ZF domain) and A3 (PCP $\alpha$  domain), in which clusters of alanine substitutions were introduced. These had no apparent effect on viral mRNA levels (Fig. 3B), while progeny virus titers were reduced by 10- and 200-fold, respectively (Fig. 3C), in supernatants harvested 24 h after transfection, well beyond the time point of maximum virus production by the



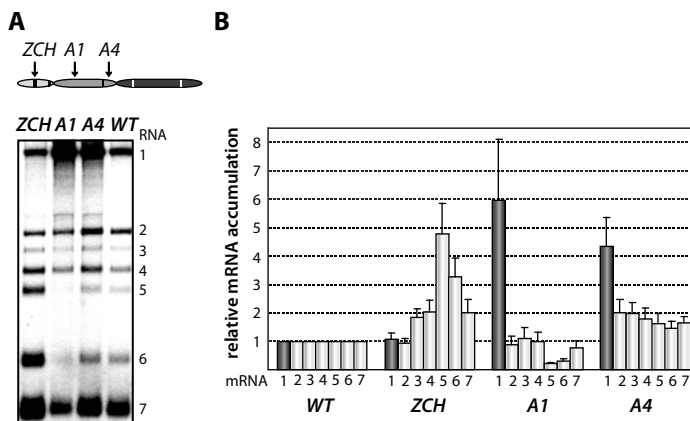
**Figure 3.** Importance of nsp1 subdomains for transcription and virus production. (A, B). Analysis of EAV-specific mRNA accumulation by gel hybridization. The domain organization of nsp1 is depicted as in Fig. 2 and the positions of the clusters of amino acid mutations analyzed are indicated with arrows. BHK-21 cells were transfected with RNA transcribed from wt or selected mutant EAV full-length cDNA clones. Total intracellular RNA was isolated at 11 h post-transfection and resolved by denaturing formaldehyde electrophoresis. Equal loading of samples was confirmed by ethidium bromide staining of ribosomal RNA (data not shown). EAV-specific mRNAs were detected by hybridization of the gel with a <sup>32</sup>P-labelled probe complementary to the 3' end of the viral genome and subsequent phosphorimaging. The positions of the EAV genome (RNA1) and the six sg mRNAs (RNA2 to RNA7) are indicated. (C) Plaque phenotype and virus titers of the Z2 and A3 mutants. Plaque assays were performed on BHK-21 using cell culture supernatants harvested 24 h after transfection. Cells were incubated under a semi-solid overlay at 39.5°C for 72 h, fixed and stained with crystal violet. Virus titers represent an average of three independent experiments. pfu, plaque-forming units.



wt control (data not shown). Accordingly, plaques of the Z2 mutant were somewhat smaller than those of the wt virus, and those of the A3 mutant were minute (Fig. 3C). Titers and plaque phenotypes remained essentially unchanged at 48 h post-transfection, arguing against a delay in virus production. Sequence analysis of the A3 progeny revealed reversion of the E113A mutation to the wt sequence at later time points (data not shown). These observations imply that both the ZF and the PCPa domains of EAV nsp1 are involved in a step of the viral replicative cycle that is downstream of transcription and is critical for the efficient production of infectious virus particles.

### Mutations in nsp1 can differentially affect accumulation of viral mRNA species

Two mutant phenotypes were previously described upon examination of the role of EAV nsp1 in transcription: one in which sg mRNA accumulation was selectively abolished, and another in which the levels of all sg mRNAs were uniformly reduced relative to that of the genomic RNA<sup>37,39</sup>. In this study, the ZCH, A1, and A4 mutants displayed a third phenotype, demonstrating that replacements in nsp1 can affect EAV RNA levels in an mRNA-specific manner. The swapping of two proposed zinc-coordinating residues in the ZCH mutant resulted in the upregulation of a subset of sg mRNAs. In comparison to the wt control, the accumulation levels of RNA3, 4, 5, 6 and 7 were increased, while those of the viral genome and RNA2 remained largely unchanged (Fig. 4, A and B). The increase in mRNA levels was not uniform, being more pronounced for RNA5 and RNA6 (4.5-fold and 3-fold, respectively) than for RNAs 3, 4, and 7 (~2-fold). Also, the substitution of two positively



**Figure 4.** Multiple mutations in nsp1 exert species-specific effects on viral mRNA accumulation. (A, B) Gel hybridization analysis and quantification of EAV-specific mRNA accumulation in cells transfected with the ZCH, A1, A4 mutant or a wt control. (A) Viral mRNA accumulation was analyzed at 11 h post-transfection by gel hybridization as described in the legend to Fig. 3. (B) The accumulation levels of each viral mRNA in the nsp1 mutants were quantified by phosphorimaging in the linear range of exposure and normalized to the level of accumulation of each corresponding viral mRNA in the wt control, which was set at 1. Genomic RNA levels are represented as dark grey bars. The relative values correspond to the means from three independent transfections and error bars denote standard deviation.

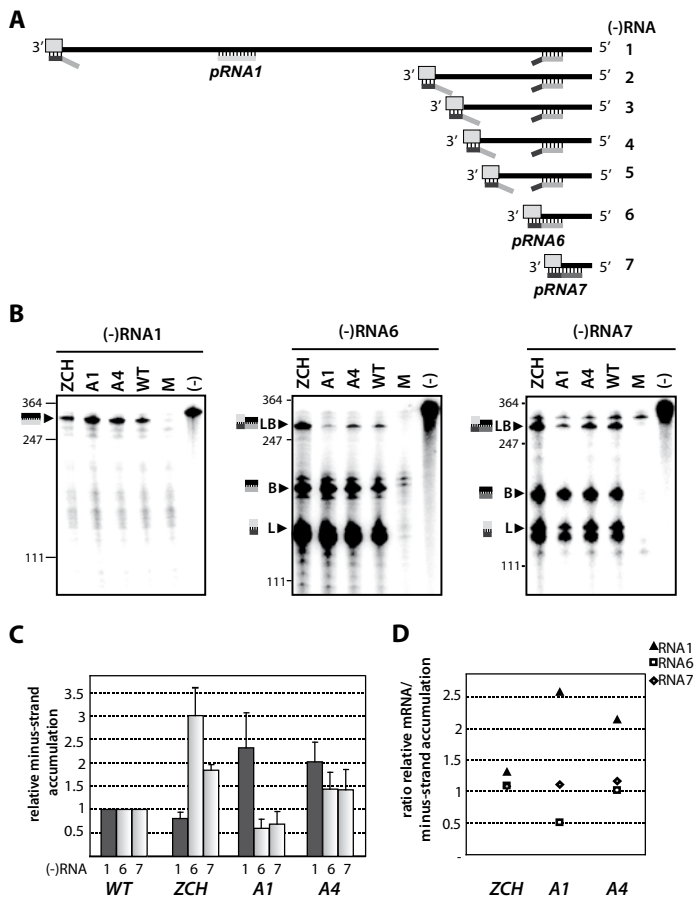
charged residues in the PCP $\alpha$  domain of the A1 mutant (see Fig.2) resulted in reduced accumulation levels of RNA5 and RNA6 (3- to 4-fold), and RNA7 (~30%), but not of RNAs 2 to 4 (Fig. 4, A and B). In contrast, genomic RNA accumulation was dramatically enhanced in the A1 mutant (Fig. 4, A and B). This aspect of the mutant phenotype had been previously described in mutants that did not produce any sg mRNAs<sup>37</sup>, in which it was attributed to the increased availability of key factors for viral replication. This explanation does not seem likely for the A1 mutant, however, in which accumulation of only two of the six viral sg mRNAs was reduced (Fig. 4). Furthermore, another PCP $\alpha$  mutant - A4 (see Fig. 2), displayed 4-fold higher levels of genome accumulation without any significant decrease in sg mRNA production (Fig. 4, A and B).

Since EAV mRNAs accumulate to different levels in molar ratios that are maintained through most of the replicative cycle (see Fig. 1B), we will refer to these levels in the wt control as “balanced”. By contrast, the ZCH, A1, and A4 mutants exerted differential effects on viral mRNA abundance, resulting in “imbalanced” viral mRNA accumulation profiles. The unusual phenotypes of these mutants implicated nsp1 in the mRNA-specific modulation of viral RNA levels, and prompted us to investigate their molecular basis in greater detail.

### **Accumulation of viral mRNAs and their corresponding minus-strand RNA templates are similarly affected by nsp1 mutations**

According to the current consensus in the field, minus-strand RNA synthesis in arteri- and coronaviruses can operate in either continuous or discontinuous mode, generating genome- or subgenome-length templates, respectively<sup>22,30,31</sup>. The relative abundance of the corresponding minus-strand template presumably determines the level to which each of the viral mRNAs accumulates<sup>1,29</sup>. We therefore sought to determine whether the differential effects of nsp1 mutations on the accumulation of EAV mRNA species were accompanied by changes in the levels of the corresponding minus-strand templates. To this end, we developed an RNase protection assay for the detection and quantification of EAV minus-strand RNA species. We employed a two-step protocol in which total RNA extracted from cells transfected with nsp1 mutants was first denatured and self-annealed. Due to the large excess of plus strands present in RNA samples extracted from EAV-infected cells<sup>48</sup>, all minus strands are expected to be present in duplexes after this annealing step, facilitating their subsequent reliable quantification. The remaining single-stranded RNA was then removed by RNase T1 digestion. Following inactivation of the enzyme, we added an excess of <sup>32</sup>P-labeled transcripts of positive polarity, which were derived either from a region unique to the EAV genome, or from the leader-body junction regions of RNA6 and RNA7. The samples were then subjected to a second round of RNA denaturation, hybridization, and digestion with RNase A and T1, after which the protected fragments were analyzed by electrophoresis.

The minus-strand templates of the most abundant viral mRNAs – RNA1, 6, and 7, were selected for quantitative analysis. Accumulation levels of genome-length [(-)RNA1] and



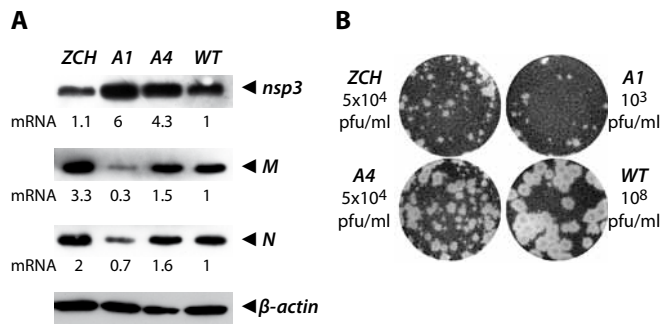
**Figure 5.** Minus-strand RNA accumulation is also modulated by mutations in nsp1. (A-D) Analysis and quantification of EAV minus-strand accumulation by a two-cycle RNase protection assay. (A) Schematic representation of the nested set of viral minus-strand RNA [(-)RNA] species produced in EAV-infected cells. The anti-leader sequence is depicted as a grey box. The *in vitro*-transcribed plus-strand probes used for detection of (-)RNA1 (pRNA1), (-)RNA6 (pRNA6) and (-)RNA7 (pRNA7) are shown. pRNA6 and pRNA7 target the leader-body junction sequences of (-)RNA6 and (-)RNA7, respectively. Note that hybridization with pRNA1 results in the protection of a single fragment, while the probes for (-)RNAs 6 and 7 each protect three fragments – one derived from the full-length sg minus strand, and two fragments derived in part from partial hybridization of these probes to larger viral (-)RNAs in which the target sequences are noncontiguous (exemplified for pRNA6). For simplicity, non-EAV sequences present near the termini of the three probes were omitted from the scheme. (B) Viral (-)RNA accumulation was analyzed at 11 h post-transfection for the ZCH, A1 and A4 mutants, and a wt control. Protected fragments were resolved on denaturing 5% polyacrylamide/8M urea gels and visualized by phosphorimaging. The constructs analyzed are labeled above the lanes (M, mock-transfected cells; (-), no-RNase control that shows a band corresponding to 0.2 fmol of the full-length probe). Sizes (nt) of RNA markers have been indicated on the left. The single 327-nt protected fragment resulting from hybridization with the positive-sense probe for RNA1(-) is indicated. The probes for subgenome-length minus strands protected fragments derived from the full-length (-)RNA6 and (-)RNA7 (327 nt and 319 nt, respectively; denoted with LB), as well as from the (-)RNA6 and (-)RNA7 body sequences (188 nt and 180 nt, respectively; denoted with B) and the anti-leader sequence (139 nt; denoted

with L). The presence of two bands in the size range of the anti-leader fragment has been described previously<sup>10</sup>. (C) The relative levels of minus-strand accumulation were quantified by phosphorimaging. For (-)RNAs 6 and 7, only the bands resulting from protection of full-length sg minus strands (denoted with LB in panel [B]) were quantified. The values correspond to the means from three independent transfections that were normalized to the level of accumulation of each minus-strand RNA in the wt control, which was set at 1. Intracellular RNA from the same transfection samples for which plus-strand accumulation was quantified (Fig. 4B) was used. Genomic minus-strand RNA levels are represented as dark grey bars. Error bars denote standard deviation. (D) The ratio of plus-strand to minus-strand accumulation for RNAs 1, 6 and 7 was calculated using the mean relative values obtained in Fig. 4B and Fig. 5C.

subgenome-length minus strands corresponding to RNA6 and RNA7 [(-)RNA6 and (-)RNA7] were quantified in total intracellular RNA extracted at 11 h post-transfection with the ZCH, A1, and A4 mutants, and a wt control. Subgenomic minus- and plus-strand levels were similarly affected by the nsp1 mutations in a sg RNA-specific manner (Fig. 5). Genomic minus-strand accumulation was increased in the A1 and A4 mutants, albeit to a somewhat lesser extent as compared to the increase in genomic plus-strands (Fig. 4). These results clearly implicate nsp1 in a regulatory step (or steps) that controls minus-strand RNA accumulation, and ultimately determines the levels to which both genome- and subgenome-length mRNA species accumulate in EAV-infected cells.

### **Changes in the accumulation of viral mRNAs induced by nsp1 mutations result in altered viral protein levels**

The relative abundance of nidovirus mRNAs likely serves to regulate the relative concentration of their respective translation products during infection. It was therefore important to determine whether viral protein levels indeed mirrored the specific changes in viral mRNA levels caused by mutations in nsp1. To this end, we examined the intracellular accumulation of the EAV replicase subunit nsp3 and the structural proteins M and N in cell lysates harvested 11 h after transfection (Fig. 6A). When compared to the wt control, nsp3 was more abundant in cells transfected with the A1 and A4 mutants, in line with the increased genome levels observed for these mutants at the same time point (Fig. 4). The intracellular levels of the M and N proteins were also in general agreement with the abundance of their corresponding mRNA templates (RNA6 and RNA7, respectively). Interestingly, even the modest reduction in RNA7 levels detected for the A1 mutant (~30%, Fig. 4B) was reflected in a decrease in N protein levels (Fig. 6A). The close correlation between mRNA and corresponding protein levels argues against the possibility that the engineered nsp1 mutations might have caused a defect in viral mRNA translation. Overall, these data establish that even modest changes in viral mRNA accumulation are directly translated into altered viral protein levels during EAV infection.



**Figure 6.** Analysis of viral protein accumulation and virus production by selected *nsp1* mutants. (A) Western blot analysis of EAV-specific protein accumulation. Cells transfected with wt or mutant viral genomes were harvested 11 h after transfection and equal amounts of total protein were analyzed with EAV-specific sera detecting *nsp3*, M and N, which are translated from RNAs 1, 6 and 7, respectively. The relative levels of each mRNA template (derived from Fig. 4B) are indicated below the gels. Beta-actin was used as a loading control. (B) Plaque phenotype and virus titers of the ZCH, A1 and A4 mutants in comparison with wt. Plaque assays were performed on BHK-21 using cell culture supernatants harvested 24 h after transfection. Virus titers represent an average of three independent experiments.

### Production of infectious virus particles is severely impaired in *nsp1* mutants with imbalanced mRNA accumulation

We next examined whether the production of infectious virus particles was disturbed in the three *nsp1* mutants with imbalanced mRNA accumulation profiles. EAV assembly involves the coordinated interplay of the N protein and six envelope proteins, all of which are required for virus infectivity<sup>20,53</sup>. A critical step in assembly is the heterodimerization of the major viral envelope proteins, GP<sub>5</sub> and M<sup>8,34</sup>, and both the GP<sub>5</sub> and M protein levels, as well as the levels of their mRNA templates, RNA5 and RNA6, were strongly affected in the ZCH and A1 mutants (Fig. 4B and 6A). Furthermore, other documented interactions, such as the oligomerization of the minor envelope proteins - GP<sub>2b</sub>, GP<sub>3</sub> and GP<sub>4</sub><sup>52</sup>, could be readily affected by the altered relative abundances of viral structural proteins. The decreased ratio of N protein and genomic RNA in the A1 and A4 mutants (Fig. 4B and Fig. 6A) could also adversely affect the assembly of infectious virions. Analysis of the infectious progeny virus titers in culture supernatants harvested 24 h after transfection with the ZCH, A1, and A4 mutants indeed revealed a dramatic loss of infectivity (Fig. 6B). In comparison to wt, the infectious progeny yield of the ZCH and A4 mutants were reduced by approximately 4 logs and that of the A1 mutant by ~5 logs.

We subsequently purified virions by sedimentation through a sucrose cushion in order to quantify their genomic RNA content by reverse transcription-quantitative PCR. Virions from medium harvested at 24 h after transfection with the ZCH, A1, and A4 mutants were compared with the progeny produced by the wt control. Consistent with the reduction in infectious progeny titers, these results showed a decrease in the total number of genome-containing virus particles secreted from cells transfected with each of the three mutants (Table 2). Interestingly, when the relative specific infectivity

of each mutant virion preparation was assessed by relating the genomic RNA content to the plaque-forming units (pfu), a marked decrease in the pfu per unit of genomic RNA ratio of the ZCH, A1, and A4 virion preparations was revealed (Table 2). The ZCH, A1, and A4 mutations thus seem to affect both the total number of secreted virions, as well as their specific infectivity. Also, even at this relatively early time point post transfection, the three mutants exhibited heterogeneous plaque morphology, indicative of rapid reversion (Fig. 6B). These results, together with the RNA and protein analyses outlined above (Fig. 4B and 6A), demonstrate that even moderate changes in the accumulation of EAV mRNAs and proteins can be associated with a dramatic decrease in the yield of infectious progeny. Thus, a previously unnoticed link seems to exist between the fine-tuning of the relative abundance of EAV mRNAs (and, consequently, viral protein levels) and the efficiency of virion biogenesis.

**Table 2.** Relative specific infectivities of virus particles from nsp1 mutants with imbalanced mRNA profiles

Construct	Relative infectivity <sup>a</sup>		Relative genomic RNA content <sup>b</sup>		Relative specific infectivity <sup>c</sup>	
	Exp. 1 <sup>d</sup>	Exp. 2	Exp. 1	Exp. 2	Exp. 1	Exp. 2
pEAV211	1	1	1	1	1	1
ZCH	$4 \times 10^{-4}$	$5 \times 10^{-4}$	$2.7 \times 10^{-1}$	$7.9 \times 10^{-1}$	$1.9 \times 10^{-3}$	$5.1 \times 10^{-4}$
A1	$2 \times 10^{-5}$	$5 \times 10^{-6}$	$8.5 \times 10^{-4}$	$6.4 \times 10^{-3}$	$2.4 \times 10^{-2}$	$7.9 \times 10^{-4}$
A4	$4 \times 10^{-4}$	$4 \times 10^{-4}$	$1.8 \times 10^{-2}$	$1.6 \times 10^{-1}$	$2.2 \times 10^{-2}$	$2.4 \times 10^{-3}$

<sup>a</sup> Virus titers were determined by plaque assays of culture supernatants harvested at 24 h post-transfection and normalized to the pfu/ml value of the wt control, which was set at 1.

<sup>b</sup> EAV genomic RNA levels were quantified in virion preparations obtained from culture supernatants harvested at 24 h post-transfection by reverse transcription and quantitative PCR. Values for mutant virions were obtained by comparing their threshold cycle (Ct) against the qPCR standard curve, and were normalized relative to the genomic RNA level in the wt control, which was set at 1.

<sup>c</sup> Relative specific infectivity values were calculated by dividing the relative infectivity (mutant:wt pfu ratio) by the relative genomic RNA content for each construct.

<sup>d</sup> The data shown are derived from two independent experiments.

## Second-site mutations in nsp1 can restore both the quantitative balance among viral mRNA species and efficient virus production

To gain more insight into the molecular basis of the nsp1 mutant phenotypes described above, we attempted to isolate revertant viruses encoding compensatory second-site mutations. The mutants in which sg mRNA accumulation was completely blocked, however, proved to be extremely stable. We were repeatedly unable to detect infectious particles in supernatants of transfected cells, even after prolonged incubation (up to 70 h) at 39.5°C or at a reduced temperature of 35°C, or after using these supernatants to infect fresh BHK-21 cells (data not shown). By contrast, the rapid appearance of large plaque variants among the prevailing small plaques produced by the ZCH, A1, and A4 mutants (see Fig. 6B) suggested genetic heterogeneity. Large plaque clones of these three mu-

tants were isolated and propagated in fresh cells, and the EAV nsp1 gene was amplified by RT-PCR. Sequence analysis of the PCR products confirmed the presence of the originally engineered mutant codons, and identified additional mutations in the nsp1-coding sequence in the majority of the plaques analyzed (data not shown). A pseudorevertant of ZCH had acquired a substitution in the vicinity of the original mutations (Ala-29 to Asp). Interestingly, Ala-29 was also mutated in five independent A1 pseudorevertants, where it had been replaced with Lys due to two nucleotide substitutions (Table 3). In addition, four clones of the A4 offspring contained a Thr-196 to Lys substitution, and Gly-47 to Ala and Glu-112 to Lys replacements were found in one clone each.

**Table 3.** Overview of EAV nsp1 pseudorevertants described in this study.

Construct	Second-site mutation <sup>a</sup>	Wild-type codon <sup>b</sup>	Mutant codon	Titer (pfu/ml) <sup>c</sup>	Plaque size
pEAV211	NA <sup>d</sup>	NA	NA	1×10 <sup>8</sup>	wild-type
ZCH	A29D (1)	309-GCC <sup>e</sup>	GAC	5×10 <sup>6</sup>	intermediate
A1	A29K (5)	309-GCG	AAG	5×10 <sup>6</sup>	intermediate
A4	G47A (1)	363-GGU	GCU	5×10 <sup>7</sup>	wild-type
	E112K (1)	558-GAA	AAA	5×10 <sup>7</sup>	wild-type
	T196K (4)	810-ACA	AAA	5×10 <sup>7</sup>	wild-type

<sup>a</sup> Virus clones were isolated from plaque assays of culture supernatants harvested between 20 h and 24 h post-transfection. The number of clones containing each mutation is indicated in brackets.

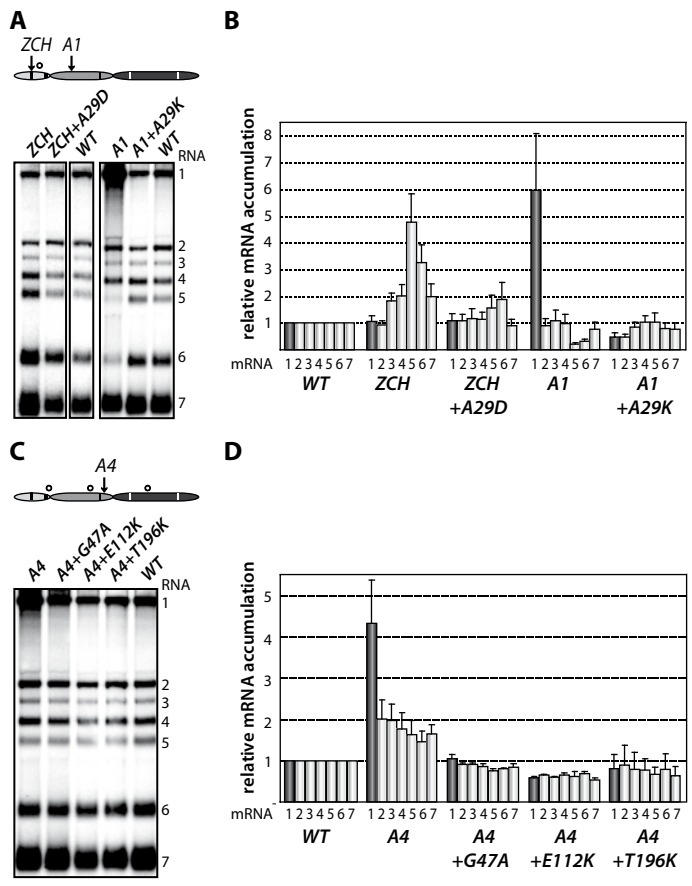
<sup>b</sup> Numbers indicate the start coordinate of the codon in the EAV genome.

<sup>c</sup> Pseudorevertants were reconstructed in the wild-type EAV background, and infectious progeny titers were determined by plaque assays of culture supernatants harvested at 24 h post-transfection. The average titers of three independent experiments are shown.

<sup>d</sup> NA, not applicable.

<sup>e</sup> A silent mutation changing the original alanine codon from GCG to GCC was introduced as a marker mutation upon construction of the ZCH mutant.

To ascertain these second-site substitutions conferred a replicative advantage, they were introduced into their respective parental (mutant) full-length cDNA clones, yielding a set of viruses collectively referred to as “nsp1 pseudorevertants”. Cell culture supernatants harvested 24 h after transfection with RNA from nsp1 pseudorevertants indeed contained between 2 and ~4 log units more infectious progeny than those of the original mutants (Table 3), confirming the compensatory nature of the second-site mutations. Both virus titer and plaque size of the three viruses carrying a second-site mutation in the A4 mutant background were similar to those of the wt control (Table 3; data not shown). Replacement of Ala-29 with Asp or Lys in the ZCH and A4 backgrounds, respectively, increased virus titers by 2 to ~4 log units, with plaque sizes being intermediate between those of the parental mutant and the wt control (Table 3; data not shown). The relative specific infectivities of virion preparations derived from all nsp1 pseudorevertants were also considerably higher in comparison to those of the parental mutants (data not shown). Notably, the pseudorevertants showed a partial or complete restoration of the mRNA-specific defects that we had observed for the original mutants (Fig. 7). Introduc-



**Figure 7.** Second-site mutations in nsp1 moderate species-specific defects in mRNA accumulation. (A-D) Gel hybridization analysis and quantification of EAV-specific mRNA accumulation. (A) BHK-21 cells were transfected with ZCH and A1 mutants, reconstructed pseudorevertants, and wt controls. The positions of the originally mutated amino acid clusters are indicated with arrows; the open circle denotes the position of second-site mutations. Viral mRNAs were analyzed 11 h post-transfection by gel hybridization as described above. (B) For ZCH+A29D and A1+A29K, the accumulation levels of each viral mRNA were quantified by phosphorimaging and the values were normalized to the wt level of accumulation of each corresponding viral mRNA from the same experiment, set at 1. Genomic RNA levels are shown as dark grey bars. The relative values correspond to the means from three independent transfections and error bars denote the standard deviation. The relative accumulation levels of viral mRNAs at 11 h post-transfection for the ZCH and A1 mutants are derived from Fig. 4B and are represented here to facilitate comparison between mutant and pseudorevertant phenotypes. (C) BHK-21 cells were transfected with the A4 mutant, reconstructed pseudorevertants and a wt control. The positions of the originally mutated amino acid cluster and the second-site mutations are indicated as in (A). Viral mRNAs were analyzed 11 h post-transfection. (D) For A4+G47A, A4+E112K, and A4+T196K, quantification of relative viral mRNA accumulation levels was performed as described in (B). Similarly, the relative accumulation levels of viral mRNAs at 11 h post-transfection for the A4 mutant derived from Fig. 4B are represented to facilitate comparison.



tion of the Ala-29 to Asp substitution in the ZCH mutant was accompanied by a considerable reduction of the otherwise abnormally high accumulation of RNAs 3 to 7, though RNA5 and RNA6 were still present at ~150% of the normal level (Fig. 7B). Likewise, the Ala-29 to Lys replacement moderated the effect of the A4 substitutions on the accumulation levels of RNA1, RNA5, and RNA6; those of RNA1 and RNA2 were somewhat reduced relative to wt. Introduction of G47A, E112K, or T196K in the A4 background in each case suppressed the increased ratio of genomic to sg mRNA that was characteristic for the parental mutant virus (Fig. 7D). Thus, both the pronounced mRNA-specific accumulation defects, as well as the associated drop in virus production observed for the ZCH, A1, and A4 mutants (Fig. 4) were considerably alleviated by second-site mutations in nsp1. The location of both the original and the second-site mutations in nsp1 revealed multiple genetic interactions between all nsp1 subdomains that are important for the protein's role in regulating the relative abundance of EAV mRNAs. In addition, the increased virus production by the pseudorevertants was invariably correlated with a distinct shift towards an mRNA accumulation profile that was (more) similar to wt (Fig. 7, B and D). These observations provide compelling evidence that efficient virus production depends on maintaining "balanced" viral mRNA levels in EAV-infected cells.

## DISCUSSION

In order to progress through their replicative cycle, +RNA viruses need to orchestrate key macromolecular processes that often overlap in time and possibly also in space. Multifunctional viral proteins, which can regulate and interlink multiple of these steps, may therefore have evolved in order to facilitate this spatio-temporal coordination. EAV nsp1 was first identified as a general transcription factor that is also critical for replicase polyprotein processing and efficient production of infectious progeny<sup>37,39</sup>. By employing forward and reverse genetics, we have now examined the significance of all nsp1 subdomains for these multiple regulatory roles. Detailed quantitative analyses of viral mRNA species and their corresponding minus-strand templates in nsp1 mutants allowed us to identify a previously unrecognized function of the protein in the mRNA-specific regulation of viral RNA abundance. Nsp1 seems to perform this function by controlling the accumulation of the minus-strand templates for viral mRNA synthesis. Together with its well-documented general role in transcription, these results place nsp1 at the heart of the unique process of discontinuous minus-strand RNA synthesis that is employed by arteriviruses and other nidoviruses. Even the modest changes in viral mRNA levels we observed in nsp1 mutants with differential deviations in viral mRNA abundance are seemingly reflected in altered levels of their translation products. The dramatically reduced virus yields of these mutants revealed that the critical fine-tuning of the "balance" of EAV mRNA accumulation has major implications for the final stages of the replicative cycle. The importance of this balance is greatly emphasized by the rapid emergence of pseudorevertant viruses, in which both the mRNA-specific accumulation defects and the

impaired virus production were significantly moderated. The mapping of compensatory second-site mutations to the nsp1 gene itself implies that multiple physical interactions among nsp1 subdomains are essential for the protein's role in transcriptional control. Our findings constitute the first evidence for the involvement of a protein factor in regulating the relative abundance of individual mRNA species during nidovirus infection.

### **The intricate interplay between nsp1's subdomains**

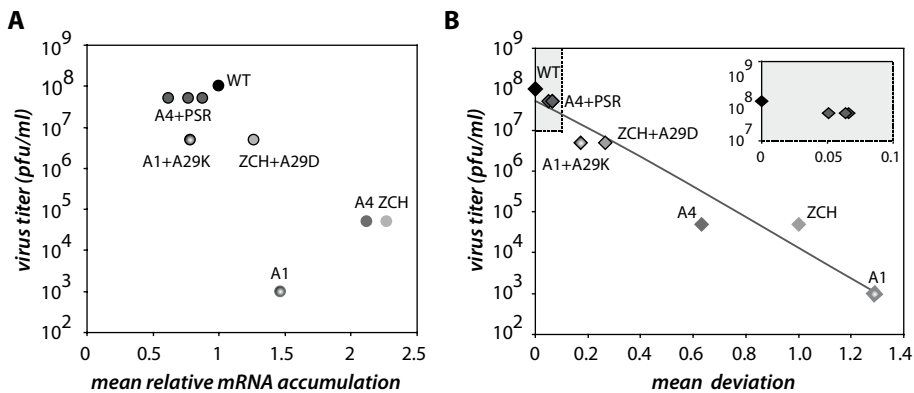
Previous reports have suggested that the multiple roles of nsp1 during EAV infection can (in part) be functionally separated, with the ZF domain being essential for transcription and efficient virus production, and the PCP $\beta$  protease cleaving the nsp1/2 site, irrespective of the ZF integrity<sup>37</sup>. This study extended and refined the above concept showing, first, that the functional repertoire of nsp1 also includes the differential control of mRNA accumulation, and, second, that some functions may be based on the interplay between two or even all three of the protein's subdomains. The phenotypes observed upon replacing charged amino acid clusters with alanine established the importance of all nsp1 subdomains for sg mRNA accumulation (Fig. 3A). The involvement of PCP $\beta$  in transcription is seemingly unrelated to its autoproteolytic function, whose inactivation is detrimental to genome replication<sup>37</sup>. The co-translational, *cis*-cleavage of the nsp1/2 site in the nascent replicase polyproteins<sup>35</sup> is probably the sole processing step mediated by PCP $\beta$ , which becomes rapidly available to exercise any *trans*-acting, non-proteolytic activities it may have. Such secondary non-proteolytic functions have previously been reported for several +RNA viral autoprotease domains, such as those of hepatitis C virus NS2<sup>14</sup> and beet yellow virus L-Pro<sup>26</sup>.

The previously uncharacterized PCP $\alpha$  domain, a PCP $\beta$  paralog that has lost its proteolytic capacity in EAV<sup>9</sup>, appears to cooperate with the ZF in transcription and virion biogenesis (Fig. 3). Likewise, it works in concert with both flanking nsp1 subdomains in controlling the relative abundance of viral mRNAs (Fig. 4, Table 3, and Fig. 7). The Lys and Arg residues replaced with alanine in the A1 and A4 mutants, and the Lys that had evolved in their pseudorevertants are all basic residues. They are found in different nsp1 subdomains, but could well be spatially juxtaposed to provide a positively charged side chain to a functional region that is involved in interactions essential for one or more of the protein's activities. The Gly-47 to Ala reversion, which maps to the junction region between the ZF and PCP $\alpha$ , might serve to reposition these subdomains relative to each other. A similar readjustment of the protein's tertiary structure might account for the compensatory effect exerted by the Ala-29 to Asp reversion on the ZCH mutant. Despite the fact that its orthologs in other arteriviruses have retained their proteolytic activity, it is tempting to speculate that the incorporation of the PCP $\alpha$  domain in the arterivirus nsp1 region can be primarily attributed to the non-proteolytic functions outlined above. On the whole, there seems to be considerable cooperation between nsp1 subdomains. This interplay appears crucial for coupling the different processes that nsp1 controls and may be based on interactions that are either intra- or intermolecular, in view of the protein's

ability to form homo-oligomers<sup>38</sup>. Notably, a recent paper describing the crystal structure of PRRSV nsp1 $\alpha$ , an arterivirus ortholog of the EAV ZF and PCP $\alpha$  domains, reported that the protein exists in equilibrium between monomers and dimers in solution. Residues from both subdomains also contribute both to nsp1 $\alpha$  dimerization and the formation of a hydrophilic groove at the dimer surface<sup>36</sup>.

### **The quantitative balance among EAV mRNA species is critical for efficient virus production**

Characterization of the three nsp1 mutants with imbalanced viral mRNA accumulation profiles showed that the disruption of the balance was due to a reduction of the levels of certain viral mRNA species only for the A1 mutant (see Fig. 4B). By contrast, specific upregulation of most mRNAs was observed for both the ZCH and A4 mutants and the associated dramatic defects in virus production were surprising, also in view of the apparently undisturbed translation of viral mRNAs (Fig. 6A). In an attempt to quantify the relationship between virus yield and viral mRNA accumulation, we used the data sets of Fig. 4B, 7B, and 7D to calculate the mean relative mRNA accumulation for the ZCH, A1, and A4 mutants, as well as their pseudorevertants. Plotting these values versus the corresponding infectivity titers (derived from Fig. 6B and Table 3) revealed efficient production of infectious progeny for viruses with mRNA accumulation that was close to that of the wt control or modestly decreased (Fig. 8A). In contrast, severely reduced virus yields were observed when viral mRNA accumulation was increased, as seen for the ZCH, A1, and A4 mutants. This somewhat counterintuitive observation was rationalized by assessing the importance of the quantitative balance among EAV mRNA species for infectious virus production. In order to establish this relationship, for each mRNA species, the absolute deviation of its relative accumulation from the mean of the complete nested set of mRNAs (Fig. 8A) was calculated. From these seven values, the mean (absolute) deviation was calculated for each mutant or pseudorevertant, and these values were also plotted against virus titers (Fig. 8B). By definition, the mean deviation is equal to zero for the wt virus, whose mRNA levels we refer to as “balanced”. All nsp1 mutants and their pseudorevertants have mean deviations larger than zero and these values reflect the magnitude of imbalance of their mRNA accumulation profiles. Remarkably, the plot revealed that the data nicely fit ( $R^2=0.95$ ) a negative exponential regression between infectious virus yield and mRNA imbalance, which is depicted as a negative linear regression in the semi-logarithmic plot of Fig. 8B. This strong relationship underlines that the magnitude of imbalance between different mRNA species, rather than the accumulation levels of mRNA species *per se*, is a chief factor affecting progeny yield. It should be noted that several factors may have affected our analysis of this relationship to a certain extent. For example, only a relatively small variety of nsp1 mutants and pseudorevertants was analyzed and although some pseudorevertants were recovered repeatedly, they were plotted only once. Also, the rapid emergence of pseudorevertants likely contributed to the virus titers measured for the ZCH, A1, and A4 mutants a 24 h post transfection, which would thus be



**Figure 8.** Relationship between viral mRNA accumulation profiles and infectious virus yield. The data obtained on the accumulation of genome and sg mRNAs were used for a quantitative assessment of nsp1 mutant phenotypes in terms of (A) changes in mRNA accumulation compared to wt and (B) the extent to which the balance between viral transcripts was disturbed (B). (A) For each mutant and pseudorevertant, a value for the accumulation of each of its 7 mRNAs at 11 h post-transfection had been assigned as compared to that of the wt control (see Fig. 4B, 7B and 7D). The mean of these seven values (“mean relative mRNA accumulation”) was plotted against the corresponding progeny virus titer in culture supernatants at 24 h post-transfection (Fig. 6B and table 2). Wild-type is depicted in black. The engineered nsp1 mutants and nsp1 pseudorevertants of mutant A4, displaying very similar phenotypes, are designated as A4+PSR. (B) For each mRNA species, the deviation of its relative accumulation from the mean of the complete nested set of mRNAs (see panel A) was calculated. From these seven values, the mean (absolute) deviation was calculated for each mutant and plotted against virus titers as in (A). For wt, the mean deviation is 0. Engineered nsp1 mutants and their pseudorevertants are indicated as in (A). The data fit a negative exponential regression calculated using Microsoft Excel and depicted as a gray line ( $y = 6 \times 10^9 + 7e^{-8.4306x}$ ,  $R^2=0.95$ ). The inset shows an expanded view of the upper left quadrant of the graph (shaded in gray).

overestimated in our analysis. Therefore, an extension of this study with new mutants and pseudorevertants may further refine the relationship illustrated by Fig. 8B. It would also be interesting to evaluate how virus yields are affected by a general imbalance between replication and transcription (A4 mutant) versus sg mRNA-specific changes (ZCH mutant), although the latter seem to have an added negative effect (A1 mutant).

The molecular basis of perturbed infectious particle production in the mutants is probably complex, although it is likely related to the altered relative abundances of viral structural proteins resulting from the respective changes in the levels of their mRNA templates (Fig. 4B and 6A). The observation that increased production of infectious progeny by the nsp1 pseudorevertants invariably correlated with restoration of balanced viral mRNA accumulation (Table 3 and Fig.7) lends further support to this hypothesis. Differential changes in structural protein levels might adversely affect their ability to form complexes that drive particle assembly, or alter the stoichiometry of these complexes when incorporated in virions. The latter option is consistent with the observed decrease in relative specific infectivity of virion preparations from the ZCH, A1, and A4 mutants (Table 2). Detailed information on the architecture of EAV virions and the molecular inter-

actions among EAV structural proteins that drive virus assembly is unfortunately lacking. Nevertheless, our data clearly establish a previously unknown link between balanced EAV mRNA accumulation and efficient virus production. We thus conclude that nsp1 critically promotes virion biogenesis by modulating viral mRNA accumulation, as well as acting at an additional, currently unknown step of the EAV replicative cycle, downstream of viral RNA synthesis<sup>37</sup> (Fig.3).

### **Nsp1 modulates a broad spectrum of successive stages in the nidovirus replicative cycle**

The onset of +RNA gene expression is marked by translation of the viral genome. In subsequent stages of infection, this molecule is utilized also as the template for replication and, in some cases, transcription, as well as packaging into progeny virus particles. Virtually nothing is known about the temporal coordination of these distinct processes in the nidovirus replicative cycle. Non-structural protein expression is extensively regulated at the translational and post-translational level, by ribosomal frame-shifting and concerted autoproteolytic processing of replicase polyproteins, which together control the production of the core viral replicative enzymes<sup>4,40,45,50</sup>. By contrast, regulation of structural protein expression is presumed to occur mainly at the transcriptional level, although the exact significance and control of the relative abundance of the various viral mRNAs in infected cells had not been examined for any nidovirus. Prior studies<sup>37,39</sup>, together with our present findings, clearly implicate EAV nsp1 in controlling the balance between replication and transcription (Fig. 3 and 4). Mutations in nsp1 were previously shown to selectively block or equally reduce the accumulation of all sg mRNA species. The pronounced upregulation of genomic RNA levels in nsp1 mutants with a complete block in sg mRNA production was also described before<sup>37</sup> and suggested to result from redirecting a limited pool of RNA-synthesizing complexes, normally engaged in both replication and transcription, to the exclusive amplification of the viral genome.

Some of the mutant phenotypes in this report, however, are poorly compatible with the above scenario. Genome RNA levels were increased 4-6 fold in the A1 and A4 mutants, for which sg mRNAs synthesis was clearly detectable and even enhanced (Fig. 4). In addition, the ZCH and A1 mutations differentially modulated the accumulation levels of a specific subset of sg mRNAs and their subgenome-length minus-strand templates (Fig. 4 and 5). This effect was accompanied by unchanged genome RNA levels in the ZCH mutant. Taken together, these data imply that nsp1 does not only allow the viral RdRp complex to engage in discontinuous minus-strand synthesis, but also enables it to differentiate between the various body TRS motifs it encounters while traversing the genomic template. This occurs by a currently unknown mechanism that is different from the previously described "polar attenuation" caused by the relative position of body TRSs in the array of successive attenuation signals that need to be "overcome" by the minus-strand RNA-synthesizing complex<sup>23</sup>. The changes in viral RNA accumulation observed for the A1 and A4 mutants could result from partial loss of recruitment of nsp1 to the RdRp complex, or its com-

promised ability to recognize RNA signals that direct discontinuous RNA synthesis. Viral RNA synthesis would then be shifted towards replication, but the increased availability of template and/or viral enzymes that this causes (Fig. 4B and 6A) might account for the relatively high levels of sg mRNA accumulation in these two mutants. This, in turn, would imply that the availability of nsp1 is important for its function in transcription. Indeed, the intracellular distribution of nsp1 is distinct from that of the other EAV nsps – a large fraction of the protein is present in the host cell nucleus<sup>38</sup>, while only ~25% of the cytoplasmic nsp1 fraction co-sediments with the membrane-bound viral RNA-synthesizing complexes upon their isolation from infected cells<sup>46</sup>. Immunofluorescence analysis did not reveal a significant change in intracellular nsp1 distribution for any of the mutants we described here (data not shown), but more rigorous biochemical studies are needed to ascertain the recruitment of nsp1 to RdRp complexes is completely unchanged. Nsp1 mutations clearly influenced minus-strand RNA accumulation (Fig. 5), although we cannot formally exclude that the protein controls minus-strand RNA stability rather than synthesis. Unfortunately, analysis of the kinetics of minus-strand accumulation in the ZCH and A1 mutants was hampered by the rapid emergence of pseudorevertants (Fig. 6B; Table 3) and the low abundance of these molecules at earlier time points post-transfection, which precluded their accurate quantitation (data not shown). Resolving this issue thus remains a formidable technical challenge. The recently described protocols for isolation of active viral RNA-synthesizing complexes from EAV-infected cells<sup>46</sup>, as well as *in vitro* activity assays using a recombinant form of the EAV RdRp<sup>2</sup> might provide better platforms for future research on the mechanistic aspects of nsp1 function.

Nsp1 remains the only known arterivirus protein specifically implicated in the regulation of transcription<sup>15,39</sup> and, to date, a functional counterpart has not been identified in coronaviruses or other nidoviruses. Species-specific changes in viral mRNA abundance were observed upon inactivation of the nsp14 exoribonuclease of human coronavirus 229E<sup>19</sup>. However, the major effect of this mutation was a reduction of the accumulation of all viral mRNAs by more than 100-fold. In view of the (indirect) dependence of transcription on genome replication and translation, the phenotype of the nsp14 mutants should be interpreted with caution. Nevertheless, a potential specific role of the coronavirus exonuclease in sg mRNA transcription deserves further investigation, although its analysis may be complicated by the multifunctionality of this protein, which was implicated in improving the fidelity of viral RNA synthesis<sup>12,33</sup> and also includes an N7-methyltransferase domain<sup>5</sup>. Elegant studies of the tombusvirus replicase have shown that genome replication and sg mRNA synthesis can be effectively uncoupled by mutations in the C-terminus of the viral RdRp, which could only be achieved after separation of the protein-coding sequence from overlapping regulatory RNA sequences<sup>54</sup>. This example once again underlines the theoretical and technical challenges encountered while dissecting the complex mechanisms coordinating +RNA virus replication and transcription.

## ACKNOWLEDGMENTS

We thank Olga Slobodskaya, Erwin van den Born, Chris Lauber, Jessika Zevenhoven and Clara Posthuma for technical assistance, sharing materials and helpful discussions.

## REFERENCES

1. **Baric, R. S. and B. Yount.** 2000. Subgenomic negative-strand RNA function during mouse hepatitis virus infection. *J.Virol.* **74**:4039-4046
2. **Beerens, N., B. Selisko, S. Ricagno, I. Imbert, Z. L. van der, E. J. Snijder, and B. Canard.** 2007. De novo initiation of RNA synthesis by the arterivirus RNA-dependent RNA polymerase. *J.Virol.* **81**:8384-8395
3. **Bordo, D. and P. Argos.** 1991. Suggestions for "safe" residue substitutions in site-directed mutagenesis. *J.Mol.Biol.* **217**:721-729
4. **Brierley, I., P. Digard, and S. C. Inglis.** 1989. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* **57**:537-547
5. **Chen, Y., H. Cai, J. Pan, N. Xiang, P. Tien, T. Ahola, and D. Guo.** 2009. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc.Natl. Acad.Sci.U.S.A* **106**:3484-3489
6. **Cunningham, B. C. and J. A. Wells.** 1989. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* **244**:1081-1085
7. **de Vries, A. A., E. D. Chirnside, M. C. Horzinek, and P. J. Rottier.** 1992. Structural proteins of equine arteritis virus. *J.Virol.* **66**:6294-6303
8. **de Vries, A. A., S. M. Post, M. J. Raamsman, M. C. Horzinek, and P. J. Rottier.** 1995. The two major envelope proteins of equine arteritis virus associate into disulfide-linked heterodimers. *J.Virol.* **69**:4668-4674
9. **den Boon, J. A., K. S. Faaberg, J. J. M. Meulenberg, A. L. M. Wassenaar, P. G. W. Plagemann, A. E. Gorbalenya, and E. J. Snijder.** 1995. Processing and evolution of the N-terminal region of the arterivirus replicase ORF1a Protein - identification of two papainlike cysteine proteases. *J. Virol.* **69**:4500-4505
10. **den Boon, J. A., M. F. Kleijnen, W. J. M. Spaan, and E. J. Snijder.** 1996. Equine arteritis virus subgenomic mRNA synthesis: analysis of leader-body junctions and replicative-form RNAs. *J.Virol.* **70**:4291-4298
11. **den Boon, J. A., W. J. M. Spaan, and E. J. Snijder.** 1995. Equine arteritis virus subgenomic RNA transcription: UV inactivation and translation inhibition studies. *Virology* **213**:364-372
12. **Eckerle, L. D., X. Lu, S. M. Sperry, L. Choi, and M. R. Denison.** 2007. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J.Virol.* **81**:12135-12144
13. **Gorbalenya, A. E., L. Enjuanes, J. Ziebuhr, and E. J. Snijder.** 2006. Nidovirales: evolving the largest RNA virus genome. *Virus Res.* **117**:17-37
14. **Jones, C. T., C. L. Murray, D. K. Eastman, J. Tassello, and C. M. Rice.** 2007. Hepatitis C virus p7 and NS2 proteins are essential for production of infectious virus. *J.Virol.* **81**:8374-8383
15. **Kroese, M. V., J. C. Zevenhoven-Dobbe, J. N. Bos-de Ruijter, B. P. Peeters, J. J. Meulenberg, L. A. Cornelissen, and E. J. Snijder.** 2008. The nsp1alpha and nsp1 papain-like autoproteases are essential for porcine reproductive and respiratory syndrome virus RNA synthesis. *J.Gen.Virol.* **89**:494-499



16. **Landt, O., H. P. Grunert, and U. Hahn.** 1990. A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene* **96**:125-128
17. **MacLachlan, N. J., U. B. Balasuriya, J. F. Hedges, T. M. Schweidler, W. H. McCollum, P. J. Timoney, P. J. Hullinger, and J. F. Patton.** 1998. Serologic response of horses to the structural proteins of equine arteritis virus. *J.Vet.Diagn.Invest* **10**:229-236
18. **Miller, W. A. and G. Koev.** 2000. Synthesis of subgenomic RNAs by positive-strand RNA viruses. *Virology* **273**:1-8
19. **Minskaia, E., T. Hertzig, A. E. Gorbalenya, V. Campanacci, C. Cambillau, B. Canard, and J. Ziebuhr.** 2006. Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc.Natl.Acad.Sci.U.S.A* **103**:5108-5113
20. **Molenkamp, R., H. van Tol, B. C. D. Rozier, Y. van der Meer, W. J. M. Spaan, and E. J. Snijder.** 2000. The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription. *J.Gen.Virol.* **81**:2491-2496
21. **Novak, J. E. and K. Kirkegaard.** 1991. Improved method for detecting poliovirus negative strands used to demonstrate specificity of positive-strand encapsidation and the ratio of positive to negative strands in infected cells. *J.Virol.* **65**:3384-3387
22. **Pasternak, A. O., W. J. Spaan, and E. J. Snijder.** 2006. Nidovirus transcription: how to make sense...? *J.Gen.Virol.* **87**:1403-1421
23. **Pasternak, A. O., W. J. M. Spaan, and E. J. Snijder.** 2004. Regulation of relative abundance of Arterivirus subgenomic mRNAs. *J. Virol.* **78**:8102-8113
24. **Pasternak, A. O., E. van den Born, W. J. M. Spaan, and E. J. Snijder.** 2001. Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. *EMBO J.* **20**:7220-7228
25. **Pedersen, K. W., Y. van der Meer, N. Roos, and E. J. Snijder.** 1999. 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.* **73**:2016-2026
26. **Peng, C. W., A. J. Napuli, and V. V. Dolja.** 2003. Leader proteinase of beet yellows virus functions in long-distance transport. *J.Virol.* **77**:2843-2849
27. **Perlman, S. and J. Netland.** 2009. Coronaviruses post-SARS: update on replication and pathogenesis. *Nat.Rev.Microbiol.* **7**:439-450
28. **Posthuma, C. C., K. W. Pedersen, Z. Lu, R. G. Joosten, N. Roos, J. C. Zevenhoven-Dobbe, and E. J. Snijder.** 2008. Formation of the arterivirus replication/transcription complex: a key role for nonstructural protein 3 in the remodeling of intracellular membranes. *J.Virol.* **82**:4480-4491
29. **Sawicki, D., T. Wang, and S. Sawicki.** 2001. The RNA structures engaged in replication and transcription of the A59 strain of mouse hepatitis virus. *J.Gen.Virol.* **82**:385-396
30. **Sawicki, S. G. and D. L. Sawicki.** 1995. Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. *Adv.Exp.Med.Biol.* **380**:499-506
31. **Sawicki, S. G., D. L. Sawicki, and S. G. Siddell.** 2007. A contemporary view of coronavirus transcription. *J.Virol.* **81**:20-29
32. **Scheltinga, S. A., K. E. Templeton, M. F. Beersma, and E. C. Claas.** 2005. Diagnosis of human metapneumovirus and rhinovirus in patients with respiratory tract infections by an internally controlled multiplex real-time RNA PCR. *J.Clin.Virol.* **33**:306-311
33. **Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. M. Poon, Y. Guan, M. Rozanov, W. J. M. Spaan, and A. E. Gorbalenya.** 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J. Mol. Biol.* **331**:991-1004



34. **Snijder, E. J., J. C. Dobbe, and W. J. M. Spaan.** 2003. Heterodimerization of the two major envelope proteins is essential for arterivirus infectivity. *J. Virol.* **77**:97-104
35. **Snijder, E. J., A. L. M. Wassenaar, and W. J. M. Spaan.** 1992. The 5' end of the equine arteritis virus replicase gene encodes a papain-like cysteine protease. *J. Virol.* **66**:7040-7048
36. **Sun, Y., F. Xue, Y. Guo, M. Ma, N. Hao, X. C. Zhang, Z. Lou, X. Li, and Z. Rao.** 2009. Crystal structure of porcine reproductive and respiratory syndrome virus (PRRSV) leader protease nsp1a. *J. Virol.* **83**:10931-10940
37. **Tijms, M. A., D. D. Nedialkova, J. C. Zevenhoven-Dobbe, A. E. Gorbalenya, and E. J. Snijder.** 2007. Arterivirus subgenomic mRNA synthesis and virion biogenesis depend on the multifunctional nsp1 autoprotease. *J. Virol.* **81**:10496-10505
38. **Tijms, M. A., Y. van der Meer, and E. J. Snijder.** 2002. Nuclear localization of non-structural protein 1 and nucleocapsid protein of equine arteritis virus. *J. Gen. Virol.* **83**:795-800
39. **Tijms, M. A., L. C. van Dinten, A. E. Gorbalenya, and E. J. Snijder.** 2001. A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus. *Proc. Natl. Acad. Sci. U.S.A.* **98**:1889-1894
40. **van Aken, D., J. Zevenhoven-Dobbe, A. E. Gorbalenya, and E. J. Snijder.** 2006. Proteolytic maturation of replicase polyprotein pp1a by the nsp4 main proteinase is essential for equine arteritis virus replication and includes internal cleavage of nsp7. *J. Gen. Virol.* **87**:3473-3482
41. **van den Born, E., A. P. Gultyaev, and E. J. Snijder.** 2004. Secondary structure and function of the 5'-proximal region of the equine arteritis virus RNA genome. *RNA* **10**:424-437
42. **van den Born, E., C. C. Posthuma, A. P. Gultyaev, and E. J. Snijder.** 2005. Discontinuous subgenomic RNA synthesis in arteriviruses is guided by an RNA hairpin structure located in the genomic leader region. *J. Virol.* **79**:6312-6324
43. **van der Meer, Y., H. van Tol, J. K. Locker, and E. J. Snijder.** 1998. ORF1a-encoded replicase subunits are involved in the membrane association of the arterivirus replication complex. *J. Virol.* **72**:6689-6698
44. **van Dinten, L. C., J. A. den Boon, A. L. M. Wassenaar, W. J. M. Spaan, and E. J. Snijder.** 1997. An infectious arterivirus cDNA clone: Identification of a replicase point mutation that abolishes discontinuous mRNA transcription. *Proc. Natl. Acad. Sci. U.S.A.* **94**:991-996
45. **van Dinten, L. C., S. Rensen, A. E. Gorbalenya, and E. J. Snijder.** 1999. Proteolytic processing of the open reading frame 1b-encoded part of arterivirus replicase is mediated by nsp4 serine protease and is essential for virus replication. *J. Virol.* **73**:2027-2037
46. **van Hemert, M. J., A. H. de Wilde, A. E. Gorbalenya, and E. J. Snijder.** 2008. The in vitro RNA synthesizing activity of the isolated arterivirus replication/transcription complex is dependent on a host factor. *J. Biol. Chem.* **283**:16525-16536
47. **van Marle, G., J. C. Dobbe, A. P. Gultyaev, W. Luytjes, W. J. M. Spaan, and E. J. Snijder.** 1999. Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and anti-sense transcription-regulating sequences. *Proc. Natl. Acad. Sci. U.S.A.* **96**:12056-12061
48. **van Marle, G., L. C. van Dinten, W. J. M. Spaan, W. Luytjes, and E. J. Snijder.** 1999. Characterization of an equine arteritis virus replicase mutant defective in subgenomic mRNA synthesis. *J. Virol.* **73**:5274-5281
49. **Versteeg, G. A., O. Slobodskaya, and W. J. Spaan.** 2006. Transcriptional profiling of acute cytopathic murine hepatitis virus infection in fibroblast-like cells. *J. Gen. Virol.* **87**:1961-1975
50. **Wassenaar, A. L. M., W. J. M. Spaan, A. E. Gorbalenya, and E. J. Snijder.** 1997. Alternative proteolytic processing of the arterivirus replicase ORF1a polyprotein: evidence that NSP2 acts as a cofactor for the NSP4 serine protease. *J. Virol.* **71**:9313-9322

51. **Wertman, K. F., D. G. Drubin, and D. Botstein.** 1992. Systematic mutational analysis of the yeast ACT1 gene. *Genetics* **132**:337-350
52. **Wieringa, R., A. A. de Vries, and P. J. Rottier.** 2003. Formation of disulfide-linked complexes between the three minor envelope glycoproteins (GP2b, GP3, and GP4) of equine arteritis virus. *J.Virol.* **77**:6216-6226
53. **Wieringa, R., A. A. de Vries, M. J. van der, G. J. Godeke, J. J. Onderwater, T. H. van, H. K. Koerten, A. M. Mommaas, E. J. Snijder, and P. J. Rottier.** 2004. Structural protein requirements in equine arteritis virus assembly. *J.Virol.* **78**:13019-13027
54. **Wu, B. and K. A. White.** 2007. Uncoupling RNA virus replication from transcription via the polymerase: functional and evolutionary insights. *EMBO J.* **26**:5120-5130
55. **Zuniga, S., I. Sola, S. Alonso, and L. Enjuanes.** 2004. Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *J.Virol.* **78**:980-994

# Chapter 7

## **Expression and purification of recombinant equine arteritis virus nonstructural protein 1**

Danny D. Nedialkova,  
Bruno Coutard,  
Bruno Canard, and  
Eric J. Snijder

## ABSTRACT

Arteriviruses and coronaviruses, distantly related plus-strand RNA viruses, generate an extensive nested set of subgenomic (sg) mRNAs as templates for the translation of viral structural proteins. These viral RNA molecules are copied from a complementary set of minus-strand templates that are produced via a unique mechanism of discontinuous minus-strand extension that is poorly understood. Non-structural protein 1 (nsp1) from equine arteritis virus (EAV), the prototype arterivirus, is essential for sg mRNA production but dispensable for genome replication. Apart from securing the balance between these key viral processes, nsp1 has also been shown to modulate viral RNA accumulation in an mRNA-specific manner by controlling the levels of templates required for viral mRNA synthesis. The development of *in vitro* assays to address nsp1 function, combined with solution of the protein's three-dimensional structure, could be instrumental in providing mechanistic details of the multiple roles of this protein in the viral replicative cycle. We therefore explored different expression conditions and fusion tags for the production and purification of soluble recombinant nsp1 in *Escherichia coli*. We describe the pitfalls we encountered upon purification of GST-tagged nsp1 and subsequent tag removal by proteolytic cleavage, and initial difficulties in purifying hexahistidine-tagged nsp1. We outline a successful approach for nsp1 expression from a construct that allowed its autoproteolytic release from a precursor, and which yielded soluble, stable recombinant nsp1 purified from bacterial cells that can be used for functional studies.

abstract

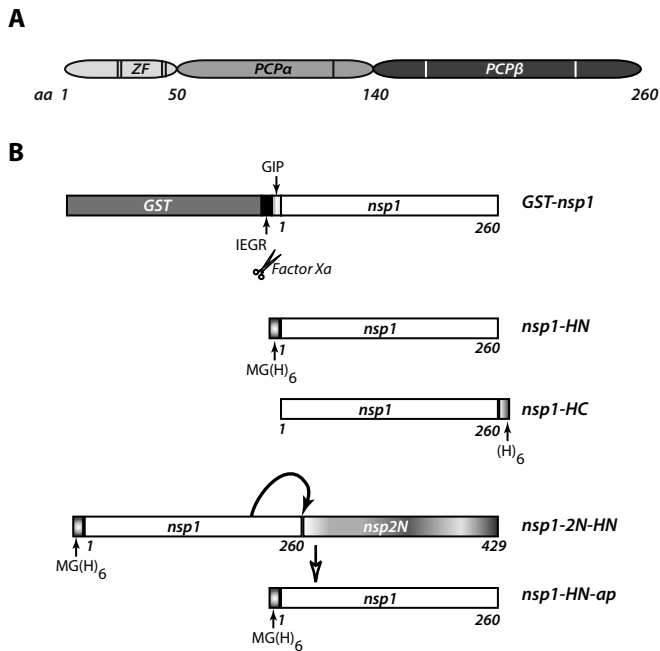
## INTRODUCTION

Replication of RNA viruses with single-stranded genomes of messenger RNA polarity (plus-strand [+]<sup>1</sup> RNA) is ensured by an RNA-dependent RNA polymerase (RdRp) that is produced by translation of the viral genome following its uncoating in the infected cell. Plus-strand RNA virus viruses also employ virally-encoded helicases (Hel), presumably to unwind local double-stranded RNA structures that might impede RdRp progress during viral RNA synthesis. The activities of these two essential enzymes need to be modulated to assure the temporal progression of the viral replicative cycle that requires the integral coordination of genome replication, translation, and packaging. +RNA viruses that produce subgenomic RNAs, like arteriviruses and coronaviruses, are challenged with an additional regulatory step. Intricate circuits involving protein-protein and protein-RNA interactions that modulate viral RNA synthesis have therefore evolved. These “molecular switches” can require the recruitment of cellular proteins, but are often mediated by viral domains that can be covalently linked to the RdRp or Hel, or exist as independent polypeptides. An example of such a candidate “molecular switch” is non-structural protein 1 (nsp1) of the equine arteritis virus (EAV), the prototype arterivirus.

The EAV genome is translated to produce a multidomain replicase precursor from the two 5'-proximal open reading frames (ORFs) in the viral genome, resulting in the synthesis of polyprotein (pp) 1a and, after a ribosomal frameshift, pp1ab. The replicase polyproteins are proteolytically processed by virus-encoded (auto)proteinases into 13 nsps that possess enzymatic activities and/or indirect regulatory roles in viral RNA replication. The viral structural genes are encoded in the 3'-end of the genome and expressed from subgenomic (sg) mRNAs. Arterivirus sg mRNAs are 5'- and 3'-coterminal and are copied from sg-length minus-strand templates produced by discontinuous minus-strand extension<sup>13</sup>, an unusual process that is also utilized by the distantly related coronaviruses and remains poorly characterized. Subgenomic mRNA synthesis and genome replication presumably compete for the same template – the genomic plus strand, implying the existence of protein and/or RNA signals that serve as regulatory switch(es).

EAV nsp1 is vital for the quantitative balance between genome replication and sg mRNA production, and is the only arteri- or coronavirus protein specifically implicated in sg mRNA synthesis to date<sup>20</sup> (Chapters 5 and 6 of this thesis). Nsp1 is also the first protein expressed during infection, by a co-translational release from the nascent replicase polyproteins by a papain-like cysteine proteinase activity (PCP $\beta$ ) in its C-terminal domain (Fig. 1A); this cleavage is essential for virus viability<sup>14</sup> (Chapter 5). Two additional conserved subdomains were identified by bioinformatics: a second, proteolytically silent PCP domain that is functional in other arteriviruses (PCP $\alpha$ ), and an N-terminal zinc finger (ZF) domain. Interestingly, nsp1 was recently found to modulate the levels at which individual viral mRNAs accumulate relative to each other during EAV infection, likely by controlling the relative abundance of their minus-strand templates (Chapter 6). Unraveling the molecular mechanisms underlying the regulatory functions of nsp1 could therefore provide invaluable insights into the poorly understood process of discontinuous

minus-strand synthesis. For that purpose, expression and purification of recombinant nsp1 and solving its three-dimensional structure would greatly facilitate its detailed characterization. The ability of the protein to form complexes with RNA and specifically recognize viral sequences that regulate sg mRNA synthesis, as well as its influence on the *in vitro* EAV RdRp activity could be examined<sup>3,22</sup>. Also, experiments based on isolated active viral RNA-synthesizing complexes from cells transfected with nsp1 mutants and addition of recombinant nsp1 could address the importance of timing of nsp1 expression and whether defects associated with mutations in nsp1 can be complemented by the wild-type protein.



**Figure 1.** Constructs for expression of recombinant EAV nsp1. (A) Domain organization of EAV nsp1. Three conserved subdomains were identified by comparative sequence analysis: an N-terminal zinc finger (ZF) domain, a papain-like cysteine proteinase activity (PCPβ) in its C-terminal domain which mediates the co-translational release of nsp1 from the nascent replicase polyproteins, and a second, proteolytically silent PCP domain that is functional in other arteriviruses (PCPa). The position of predicted zinc-coordinating residues and the active-site amino acids of PCPa and PCPβ are indicated with lines. The lack of PCPa proteolytic activity is likely due to a natural loss of the active-site Cys. (B) Schematic diagrams of fusion proteins produced from nsp1 expression constructs. Nsp1 was expressed in *E. coli* fused to the C-terminus of GST with a Factor Xa Protease cleavage site (IEGR) and a three-amino-acid spacer. Proteolytic processing of the GST-nsp1 fusion protein by Factor Xa yielded nsp1 containing three foreign amino acids (GIP) at its N-terminus. Numbers indicate amino-acid positions in EAV pp1a. Nsp1 was also produced as an N- or C-terminally His-tagged protein (nsp1-HN and nsp1-HC, respectively). A recombinant protein with an N-terminal His tag followed by the nsp1 sequence and 169 residues from nsp2 was used to generate bacterially expressed autoprocessed nsp1. This recombinant form of nsp1, whose amino acid sequence is identical to that of nsp1-HN, is referred to as nsp1-HN-ap in the text.

We therefore set out to design a protocol for the expression and purification of stable and soluble recombinant nsp1 for crystallization and structure determination via X-ray diffraction, as well as functional *in vitro* studies. Since significant amounts (mg) of protein are usually required for structural studies, and nsp1 is a relatively small protein (30 kDa) that is not expected to be post-translationally modified, we chose *E. coli* as the expression host. Several terminally extended derivatives of nsp1 were produced and a variety of conditions for their purification were explored. The nsp-coding sequence was fused to the C-terminus of glutathione-S-transferase (GST) via a linker region encoding a Factor Xa Protease cleavage site (Fig. 1B). Alternatively, hexahistidine [(His)<sub>6</sub>] tags were fused to the N-terminus or C-terminus of the protein. Lastly, expression of nsp1 with an N-terminal (His)<sub>6</sub> tag from a construct that encoded an amino-acid sequence derived from the N-terminal region of nsp2 was used to assess the effect of autoproteolytic cleavage on the solubility and stability of recombinant nsp1. We outline the technical difficulties encountered upon purification of GST-nsp1 and removal of the affinity tag by proteolysis. Furthermore, we show that autoproteolytic release of nsp1 from a precursor polypeptide is likely key for the successful purification of stable, soluble nsp1 with an N-terminal (His)<sub>6</sub> tag by metal affinity chromatography. Although the current yield of recombinant nsp1 obtained following the protocol described here is likely too low for structural studies, the purified protein can be used for setting up *in vitro* assays aimed at elucidating one or more of its functions.

## MATERIALS AND METHODS

### Construction of plasmids for recombinant nsp1 expression

In order to generate a plasmid for the expression of EAV nsp1 fused to GST in bacterial hosts, the nsp1-coding sequence (nucleotides 225 to 1004 of the wild-type EAV infectious cDNA clone pEAV211<sup>21</sup>) was amplified by PCR with oligonucleotides E872 (5'-CGCGGATCCCCATGGCAACCTTCTCCGC-3') and E386 (5'-GGAATTCGCATGCCTAGCCGTAGTTGCCAG-3'). The fragment was cloned between the *Bam*HI and *Eco*RI restriction sites of the pGEX-5X-3 expression vector (GE Healthcare, Chalfont St Giles, UK), transcription from which is driven by the chemically inducible *tac* promoter. The resulting plasmid, pGEX-5X-3-nsp1, encoded EAV nsp1 (amino acids Met-1 to Gly-260 of replicase pp1ab of the EAV-Bucyrus strain, NCBI genome database accession number NC\_002532) fused to the C-terminus of GST. The two protein moieties are separated by 7 amino acid spacer representing a Factor Xa cleavage site (NH<sub>2</sub>-Ile-Glu-Gly-Arg-COOH), followed by NH<sub>2</sub>-Gly-Ile-Pro COOH, originating from translation of sequences in the multiple cloning site of the vector.

For the generation of plasmids for bacterial expression of (His)<sub>6</sub>-tagged EAV nsp1, oligonucleotides E777 (5'-CATGCCATGGGCCATCACCATCACCATCACATGGCAACCTTCTCCGCTACTGG-3') and E386 (5'-GAATTCGCATGCCTAGCCGTAGTTGCCAG-3') were used to

184

amplify the nsp1-coding sequence, introducing a start codon followed by sequences specifying a (His)<sub>6</sub> tag, and a stop codon, respectively. Amplification with E385 (5'-TATA-ACGCGTGGTACCATGGCAACCTTCTCC-3') and E778 (5'-CCGCTCGAGTTAGTGATGGTGATG-GTGATGGCCGTAGTTGCCAGCAGGCAAAC-3') was used to introduce the (His)<sub>6</sub> tag-coding sequence followed by a stop codon, downstream of nt 1004. The fragments were cloned between the *Nco*I and *Xho*I restriction sites of Gateway entry vector pENTR11 (Invitrogen), and, after sequence verification, placed downstream of the T7 promoter in Gateway expression vector pDEST14. The resulting constructs, pDEST14-EAV nsp1-HN and pDEST14-EAV nsp1-HC, mediate the expression of recombinant nsp1 preceded by an N-terminal Met-Gly-(His)<sub>6</sub> tag (nsp1-HN), or followed by a C-terminal (His)<sub>6</sub> tag (nsp1-HC). Lastly, an expression construct encoding aa 1 to 429 of EAV pp1a preceded by a Met-Gly-(His)<sub>6</sub> tag, was generated by PCR amplification of nt 255 – 1511 of pEAV211 with E777 and E1064 (5'- CCGCTCGAGTTA TGTGGCAAGCTTGGTCCTTGG – 3'), subsequent cloning of the fragment between the *Nco*I and *Xho*I restriction sites of pENTR11 and recombination to pDEST14. The resulting plasmid, pDEST14-nsp1-2N-HN, mediates the expression of EAV nsp1 fused to 169 aa from the N terminus of EAV nsp2, allowing for the generation of autoproteolytically processed nsp1 with an N-terminal (His)<sub>6</sub> tag. This recombinant form of nsp1, whose amino acid sequence is identical to that of nsp1-HN, is referred to as nsp1-HN-ap in the text.

### **Bacterial strains for recombinant nsp1 expression**

The BL21(DE3) strain of *Escherichia coli* and its derivatives, C41 (DE3) and C43(DE3) were used as expression hosts as indicated in the text. The BL21(DE3) strain lacks the OmpT and Lon proteases and contains a chromosomal copy of the T7 polymerase gene behind a *lacUV5* promoter. Expression of T7 polymerase is induced by addition of the lactose analog isopropyl b-D thiogalactoside (IPTG). The C41 (DE3) and C43(DE3) strains were isolated in a screen for BL21(DE3) derivatives with improved characteristics of toxic protein expression<sup>12</sup>.

### **Expression and purification of GST-nsp1**

Plasmid pGEX-5X-3-nsp1 was transformed in an expression strain of *E. coli* and a single colony was used to inoculate 2 ml of LB containing 50 µg/ml of ampicillin (LB-Amp). The culture was grown overnight, diluted 1:100 in fresh LB-Amp, and subsequently grown at 37°C to mid-logarithmic phase (OD<sub>600</sub> 0.6 – 0.8). Recombinant protein expression was induced by addition of IPTG. The duration of induction, temperature at which bacteria were cultured during induction, and amount of IPTG, are indicated in the text. After induction, cells were harvested by centrifugation at 2,500 x *g* for 10 min and stored at -20°C. The pellets were resuspended in buffer A (10 mM Tris-HCl pH 7.5, 400 mM NaCl, 10% glycerol; 1 ml of buffer A was used per 10 ml of original culture volume) supplemented with 1 mg/ml lysozyme (Sigma-Aldrich) and 150 U/ml *DNase*I (Invitrogen), and incubated for 1 h



at 4°C with gentle mixing. The lysates were then supplemented with 0.5 % Triton-X100, incubated for 30 min at 4°C and clarified by high-speed centrifugation. Pellet fractions were resuspended in 1X Laemmli sample buffer (LSB) for SDS-PAGE analysis of insoluble proteins. Supernatant fractions were incubated with GST affinity resin (glutathione sepharose 4B [GS4B] from GE Healthcare, unless otherwise indicated in the text) for 1 – 2 h at 4°C with gentle mixing. The affinity resin was then washed extensively with buffer A supplemented with 0.5 % Triton-X100, followed by washing with buffer A alone. The bound GST-nsp1 was eluted with 10 mM reduced glutathione in 50 mM Tris-HCl, pH 8.0, and analyzed by SDS-PAGE and Coomassie Brilliant Blue staining. Alternatively, the resin-bound fusion protein was equilibrated in Factor Xa cleavage buffer and digested with Factor Xa Protease (see below).

### Factor Xa cleavage of GST-nsp1

GS4B resin-bound recombinant GST-nsp1 was equilibrated in FXa buffer (50 mM HEPES pH=6.5, 100 mM NaCl, 10 % glycerol, 1 mM CaCl<sub>2</sub>, 50 mM L-Arg pH=9.5; final pH of buffer: 7.5). Subsequently, Factor Xa Protease (Qiagen) was added to the slurry and proteolytic reactions were carried out as specified in the text. Cleaved nsp1, containing three additional amino acids at the N terminus (NH<sub>2</sub>- Gly-Ile-Pro COOH) as a result from the cloning strategy, was recovered from the supernatant of the pelleted GS4B resin. After digestion, Factor Xa Protease and traces of GST or uncleaved fusion protein were removed by incubation of the nsp1-containing fraction with Xa Removal Resin (Qiagen; 50 µl per 4 U of enzyme) and GS4B. The nsp-containing fraction was then centrifuged at 100,000 x g for 1 h at 4°C in a Beckman tabletop ultracentrifuge. The supernatant fraction was analyzed by gel filtration on a Superdex 75 column (GE Healthcare) in 50 mM HEPES pH=6.5, 100 mM NaCl, 10 % glycerol, 50 mM L-Arg.

### Expression and purification of His-tagged nsp1

Expression of recombinant nsp1 from pDEST14-EAV nsp1-HN and pDEST14-EAV nsp1-HC vectors was performed essentially as described above for GST-nsp1 fusion protein expression. Supernatants from induced bacterial cultures lysed in buffer B (10 mM Tris-HCl pH 8.0, 400 mM NaCl, 10% glycerol supplemented with 0.5 % Triton-X100) were incubated with Ni-NTA agarose (Qiagen) or HIS-Select HF Nickel Affinity Gel (Sigma-Aldrich) for 1 h at 4°C with gentle mixing. Unbound proteins were washed off with buffer B containing 10 mM imidazole, and proteins bound to the resin were eluted in buffer B supplemented with 250 mM imidazole.

For expression and purification of autoproteolytically processed recombinant nsp1-HN-ap, the pDEST14-nsp1-2N-HN expression vector was transformed in C41(DE3) cells and a single colony was used to inoculate 5 ml of LB-Amp. After growth for 4 h at 37°C, the starting culture was diluted in 1 l pre-warmed LB-Amp and cells were grown at 37°C until the OD<sub>600</sub> equaled 0.6. The culture was cooled on ice and recombinant protein expression

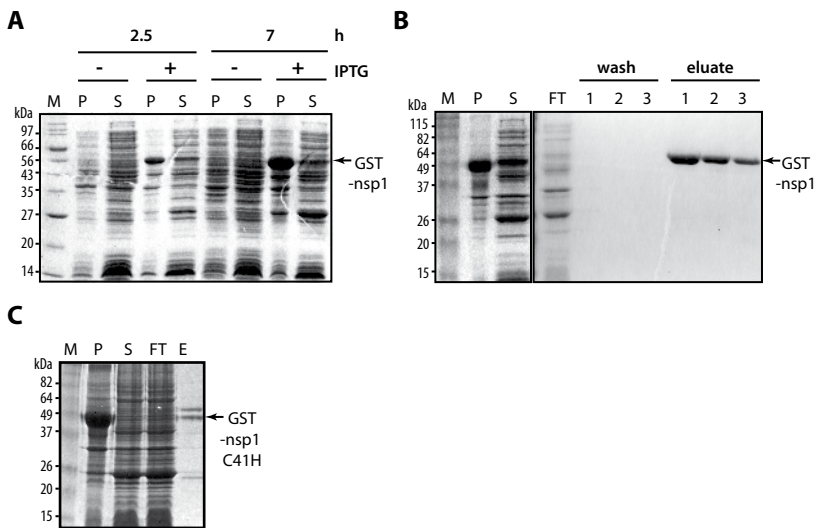
was induced with 1 mM IPTG at 16°C for 16 h. The cells were harvested by centrifugation at 2,500 x *g* for 10 min, pellets were resuspended in buffer C (50 mM 2-(*N*-morpholino)-ethanesulphonic acid [MES] pH 7.0, 500 mM NaCl, 10 % glycerol, 5 mM imidazole; 1 ml buffer per 50 ml original culture) supplemented with 1 mg/ml lysozyme, 30 U/ml DNaseI, and 0.1 % Triton-X100, and the cell suspension was incubated for 1 h at 4°C with gentle mixing. The Triton-X100 concentration was adjusted to 0.5%, lysates were incubated at 4°C for additional 30 min and subsequently clarified by centrifugation at 20,000 x *g* for 30 min. Recombinant nsp1 was purified from supernatants using Talon metal affinity resin (Clontech). The resin was washed twice with 50 resin volumes of buffer C containing 0.5 % Triton-X100, followed by three washes with 50 volumes of buffer C containing 10 mM imidazole. Elution was performed in 20 mM MES pH=7.0, 500 mM NaCl, 10 % glycerol, 150 mM imidazole in two steps. Eluates were pooled and immediately loaded onto a Superdex-75 column, and gel filtration was carried out as indicated in the text. Fractions containing nsp1 were pooled, and Ultrafree 10-kDa filter columns (Millipore) were used for protein concentration.

## RESULTS AND DISCUSSION

### Expression and purification of GST-tagged nsp1

Initial characterization of the proteolytic activity of EAV nsp1 showed that the protein is at least partially active as an autoprotease when increasingly longer N-terminal fragments of pp1a were expressed in as C-terminal GST fusions in bacterial cells<sup>14</sup>. Autoproteolytic processing suggested proper folding of nsp1 (and its PCPβ domain in particular) in a fraction of the bacterially produced fusion proteins, although both precursor and self-processed GST-containing proteins could only be recovered from the insoluble fractions of cell lysates under the experimental conditions used<sup>14</sup>. We therefore set out to explore expression and purification conditions that would allow us to produce soluble GST-nsp1, as well as protocols for the removal of the GST tag to obtain native nsp1. To that end, the nsp1-coding sequence (amino acids 1 – 260 from EAV pp1a) was cloned in frame with the GST gene in plasmid pGEX-5X-3, and the resulting plasmid was used to produce a GST-nsp1 fusion protein under the control of the strong IPTG-inducible *tac* promoter. The two protein moieties in the fusion are separated by a seven amino-acid linker containing a Factor Xa Protease cleavage site (Ile-Glu-Gly-Arg), followed by three residues (Gly-Ile-Pro) derived from an in-frame vector sequence (see Fig. 1C).

Lowering expression temperature has been shown to improve recombinant protein solubility in *E. coli*, presumably by allowing enough time for the proper folding of newly produced recombinant proteins due to slower metabolic rates<sup>10,23</sup>. Expression of GST-nsp1 in BL21(DE3) cells was therefore tested in small-scale bacterial cultures that were grown to mid-log phase, and protein production was induced at 30°C by the addition of 0.5 mM IPTG. Pellet and supernatant fractions of lysates (prepared as described in Materi-



**Figure 2.** Expression and purification of GST-nsp1. (A) BL21(DE3) cells transformed with pGEX-5X-3-nsp1 were cultured until  $OD_{600}$  of 0.7, GST-nsp1 expression was induced by the addition of 0.5 mM IPTG, and the cultures were incubated at 30°C. Lysates were obtained from non-induced or IPTG-induced cells harvested 2.5 or 7 hours post induction. After centrifugation of total cell lysates, obtained as described in Materials and Methods, the insoluble pellet (P) fractions were resuspended in 1x LSB and equal volumes of pellet and supernatant (S) fractions were separated on a 12.5 % polyacrylamide gel by SDS-PAGE. Proteins were visualized by subsequent staining of the gel with Coomassie Brilliant Blue R-250. (B) GST-nsp1 expression was induced with 0.3mM IPTG for 5.5 h at 25°C, and the soluble fraction of the total bacterial lysate was incubated with Glutathione Sepharose 4B (GS4B). The resin was pelleted by centrifugation, and unbound proteins (flow-through, FT) were collected for analysis. Impurities were then removed by washing the resin twice with buffer A with 0.5 % Triton-X100, followed by a single washing step with buffer A alone (for buffer composition, see Materials and Methods). Bound GST-nsp1 was subsequently eluted with 10 mM reduced glutathione in 50 mM Tris-HCl pH 8.0 in three steps. Proteins from the pellet (P) and soluble (S) fractions of the total bacterial lysates, as well as the FT, wash and elution steps were analyzed by SDS-PAGE and Coomassie blue staining. (C) BL21(DE3) cells were transformed with a mutant derivative of pGEX-5X-3-nsp1 containing a C41H substitution in nsp1. Expression and purification of the mutant recombinant protein was performed essentially as described in (B) but protein expression was induced at 20°C.

als and methods) from cells harvested at 2.5 h and 7 h after induction were analyzed by SDS-PAGE. Large quantities of a protein with an apparent molecular mass of 55 kDa, consistent with the expected mass of GST-nsp1, were produced exclusively in the IPTG induced bacterial cultures, and Western blotting with GST- and nsp1-specific antibodies confirmed the identity of the fusion protein (data not shown). The majority of recombinant GST-nsp1, however, was present in the insoluble fraction of cell lysates, especially after longer induction times (Fig. 2A).

In an effort to increase the fraction of soluble GST-nsp1 produced in bacterial cells, the temperature during protein expression was further decreased to 25°C, and the concentration of IPTG was lowered to 0.3 mM. These conditions resulted in a slightly higher yield of soluble fusion protein, although the bulk of GST-nsp1 still remained in insoluble aggregates (Fig. 2B). The fusion protein was purified from the soluble fraction of bacterial

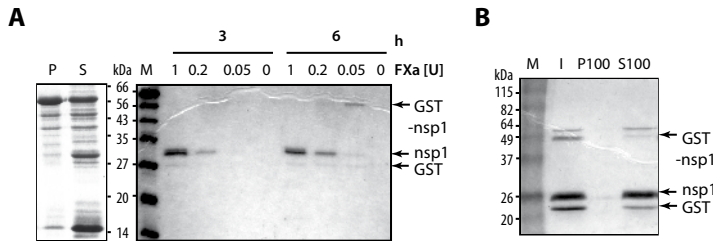
lysates with glutathione sepharose 4B (GS4B) as described in Materials and methods. The purity of the fusion protein was more than 90%, and the yield of GST-nsp1 obtained under these conditions was ~1 mg/ l culture, as estimated by Coomassie Blue staining and comparison to BSA standards (Fig. 2B and data not shown).

We then attempted to purify a mutant form of the fusion protein - GST-nsp1 C41H, carrying a Cys-41 to His substitution of a zinc-coordinating residue in nsp1 that abolishes sg mRNA production when introduced into the infectious cDNA clone of EAV (Chapter 5 of this thesis). The mutant recombinant protein, however, was found exclusively in the insoluble fraction of bacterial lysates when expressed using the protocol described above, as well as when the protein production temperature was lowered even further, to 20°C (Fig. 2C). No significant amounts of GST-nsp1 C41H were obtained upon expression of the protein in a variety of conditions aimed at promoting protein solubility and/or suppressing protein aggregation. The strategies explored included the addition of 3% ethanol to the culture medium during protein production, which was shown to mimic a heat-shock response in *E. coli* and enhance solubility of some recombinant proteins, likely by inducing overexpression of heat-shock proteins, which can act as chaperones and facilitate correct protein folding and assembly<sup>17</sup>. The increase of osmotic pressure, coupled with addition of "osmolytes" (small organic molecules, e.g. betaine and proline, which have unfavorable interactions with protein surfaces and therefore stabilize their natively folded state) to the culture medium during induction of protein expression has been reported to decrease protein aggregation in *E. coli*<sup>4,8</sup>. None of these conditions promoted greater solubility of GST-nsp1 C41H (data not shown), and neither did expression of the mutant fusion protein in two different strains of *E. coli* – C41(DE3) and C43(DE3). These results suggest the C41H substitution has a significant destabilizing effect on the overall structure of nsp1 in *E. coli*, which can be difficult to overcome, considering the predicted zinc-binding role of Cys-41 and the structural importance of zinc coordination for the proper folding of ZF domains.

### **GST tag removal by Factor Xa proteolysis**

Due to the large size of GST tag (26 kDa) and its potential to influence the biochemical properties of nsp1, tag removal was a prerequisite for subsequent functional studies and attempts to crystallize the protein in order to solve its three-dimensional structure. To this end, GST-nsp1 was expressed and purified as described above, and was subjected to proteolysis with Factor Xa protease while bound to the GST affinity resin (Fig. 3A). Addition of increasing amounts of protease resulted in the release of a 30-kDa protein in the supernatant when cleavage was performed in buffer conditions recommended by the manufacturer (Qiagen FXa buffer; see legend of Fig. 3A) at ambient temperature; this protein was absent in the no-enzyme control reaction. Western blotting confirmed that the 30-kDa cleavage product was indeed nsp1 (data not shown). Centrifugation of cleaved nsp1 at 100,000x *g*, however, revealed that the protein precipitated in the absence of the fusion tag (data not shown) in Qiagen FXa buffer. An optimized protocol for

the Factor Xa-mediated proteolytic reaction that prevented precipitation of cleaved nsp1 was then determined empirically. L-Arg, a known aggregation suppressor<sup>2</sup>, was found to be a crucial additive for stabilizing the protein (Fig. 3B), and proteolysis was carried out in FXa buffer (for composition, see the figure 3 legend) at 4° to account for potential thermal instability of nsp1. In subsequent experiments, 0.1% CHAPS was added during proteolytic digestion to decrease non specific association of cleaved nsp1 with the GS4B resin.

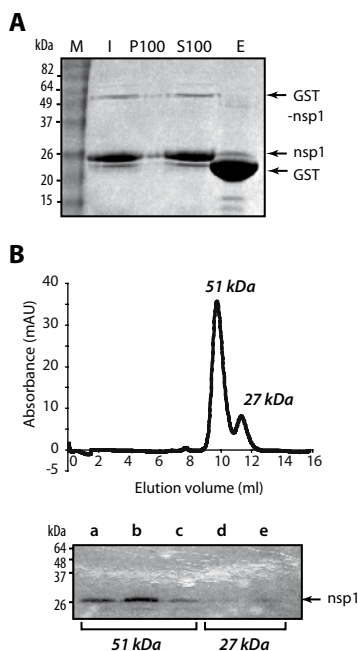


**Figure 3.** GST tag removal by Factor Xa proteolysis. (A) The soluble fraction of a total bacterial lysate (S, left panel) was divided in four samples and each sample was incubated with equal volumes of GS4B. Contaminants were washed out as described in the legend to Fig. 2B, and the resin with bound fusion protein was equilibrated in 20 mM Tris pH=6.5, 50 mM NaCl, 1 mM CaCl<sub>2</sub> (Qiagen Factor Xa buffer). Proteolytic cleavage of resin-bound GST-nsp1 was carried out at room temperature by addition of different amounts of Factor Xa Protease (1, 0.2, 0.05 or 0 units [U]). Samples were removed from the proteolytic digestion mixture at 3 and 6 hours after the start of the reaction and separated by SDS-PAGE on a 12 % polyacrylamide gel that was stained with Coomassie blue in order to estimate the yield and purity of cleaved nsp1. (B) Digestion of GS4B-bound GST-nsp1 with Factor Xa Protease was performed for 16 h at 4°C in 50 mM HEPES pH=6.5, 100 mM NaCl, 10 % glycerol, 1 mM CaCl<sub>2</sub> supplemented with 50 mM L-Arg. Cleaved nsp1 (input, I) was analyzed by ultracentrifugation at 100,000 x g for 1h at 4°C. The pellet (P100) fractions were resuspended in 1x LSB, and equal volumes of pellet and supernatant (S100) fractions were analyzed by SDS-PAGE and Coomassie blue staining as an indication of cleaved nsp1 solubility.

### Cleaved nsp1 forms dimers in solution

Using the optimized protocols for GST-nsp1 expression, purification and tag removal outlined above, a large-scale expression culture (1 l) was grown with the aim of producing sufficient recombinant nsp1 for preliminary characterization of the protein preparation's suitability for crystallization trials. Proteolysis of the resin-bound GST-nsp1 in FXa buffer resulted in the release of soluble native nsp1, and glutathione elution of GS4B-bound proteins following cleavage did not show the presence of remaining unprocessed fusion protein (Fig. 4A). Judging by the intensities of Coomassie-stained protein bands, however, the ratio of cleaved nsp1 and eluted GST protein was not 1:1, suggesting loss of native nsp1, possibly due to aggregation. In addition, impurities with molecular masses consistent with GST-nsp1 and GST remained present in the cleaved nsp1 fraction even after it was incubated with fresh GS4B resin. In an attempt to obtain a native nsp1 sample of higher purity, as well as to analyze the protein's aggregation state, the S100 fraction of cleaved nsp1 was subjected to size exclusion chromatography. The protein eluted in two peaks with estimated molecular masses of 51 kDa and 27 kDa, and SDS-PAGE analysis

of peak fractions confirmed the presence of nsp1 (Fig.4 B), suggesting that ~90% of the protein form dimers in solution under these experimental conditions. The lack of high-molecular mass peaks in the elution profile argued against the presence of non-specific aggregates in the sample. Traces of GST, however, were still found in the nsp1-containing column fractions (data not shown), likely due to the small difference in molecular mass between the two proteins and GST's ability to form homodimers<sup>9</sup>. Finally, the yield of native nsp1 obtained in this experiment was estimated to be ~0.2 mg from 1 l of culture.



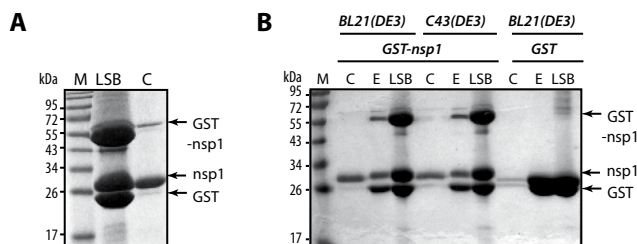
**Figure 4.** Analysis of cleaved nsp1 by gel filtration. (A) One liter of BL21(DE3) cells transformed with pGEX-5X-3-nsp1 were cultured until  $OD_{600}$  of 0.7, expression of GST-nsp1 was induced by addition of 0.3 mM IPTG and the culture was incubated for 5 h at 25°C. Cells were then harvested and lysed, and the fusion protein found in the soluble fraction of the total bacterial lysate was bound to GS4B. After washing out contaminants, 20 U of Factor Xa Protease were added, and proteolytic digestion was carried out in a 600- $\mu$ l volume for 14 h at 4°C in FXa buffer (see Materials and Methods) supplemented with 0.1 % CHAPS. Cleaved nsp1 was recovered from the resin supernatant, and GST-containing proteins remaining on the resin were eluted with 600 $\mu$ l 10 mM reduced glutathione in 50 mM Tris-HCl, pH 8.0 to check the efficiency of proteolysis (fraction E). Factor Xa and GST-containing contaminants were removed from the cleaved nsp1 fraction by subsequent incubations with Xa Removal Resin and fresh GS4B. The solubility of cleaved nsp1 was determined by ultracentrifugation as in Fig. 3B, and equal volumes of P100, S100 and E fractions were analyzed by SDS-PAGE. Half of the S100 fraction (~250  $\mu$ l) was loaded onto a Superdex-75 column. (B) Elution profile of cleaved nsp1. Gel filtration chromatography of the S100 sample in (A) was performed in 50 mM HEPES pH=6.5, 100 mM NaCl, 10 % glycerol, 50 mM L-arginine (final pH of buffer: 7.5). Blue dextran (2,000 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa) and ribonuclease A (13.7 kDa) were used as native molecular mass markers. The elution volumes of the marker proteins filtered through Superdex-75 in FXa buffer were fitted to a linear equation ( $R^2=0.98$ ), and the standard curve was used to estimate the molecular masses of nsp1 elution peaks.(C) Peak fractions from (B) were analyzed by SDS-PAGE and Coomassie Blue staining.

## Substantial protein precipitation during Factor Xa proteolysis

The production of soluble GST-nsp1 fusion in bacterial cells and nsp1 aggregation during proteolytic cleavage of the GST tag were two obvious major bottlenecks in obtaining sufficient protein material for nsp1 structure determination. Buffer composition during Factor Xa digestion of the fusion protein, however, was largely constrained by inherent properties of the protease, such as its decreased activity at high pH and salt concentrations above 100 mM. Factor Xa is also active as a disulfide-linked heterodimer, and therefore sensitive to the presence of reducing reagents during proteolytic cleavage. Addition of non ionic detergents such as Triton X-100 or Nonident P-40 reportedly does not interfere with Factor Xa proteolytic activity, although their presence even in trace amounts is poorly compatible with crystallization trials, due to their low critical micelle concentration (CMC). We therefore attempted to increase the fraction of soluble GST-nsp1 produced in expression cultures by using a number of the approaches outlined above that do not employ detergents with low CMC. A combination of expression in C43(DE3) cells and addition of 330 mM sorbitol and 2.5 mM betaine to the culture during protein production was the only condition which resulted in an increase in purified GST-nsp1, although the effects were modest at best (data not shown). In addition, a number of independent experiments for purification of GST-nsp1 from a large-scale culture and subsequent cleavage of the GST tag resulted in variable yields of cleaved nsp1 that could be as low as ~0.025 mg per liter of bacterial culture, despite high levels of GST-nsp1 in the soluble fractions of bacterial lysates.

In an effort to identify the basis of the observed variability, GST-nsp1 produced in 1 l of bacterial culture was bound to GS4B resin and subjected to Factor Xa digestion essentially as in Fig. 4. Cleaved nsp1 was recovered in the supernatant, and proteins remaining associated with the GS4B resin were eluted by boiling in an equal volume of 1x LSB. Analysis of the two fractions by SDS-PAGE revealed that by far the majority of cleaved nsp1 remained associated with the glutathione resin, as – surprisingly – did GST and a large fraction of uncleaved fusion protein (Fig. 5A). The fact that these proteins could only be dissociated from the resin by a harsh elution step is indicative of substantial protein aggregation that occurred irrespective of whether the GST-nsp1 fusion protein was produced in BL21(DE3) or C43(DE3) cells (Fig. 5B). This phenomenon could be due to the tendency of recombinant GST to undergo oxidation-induced aggregation, probably via disulfide bond formation among surface-exposed cysteine residues found close to the C-terminus of the protein<sup>1,9</sup>. Aggregated GST, however, was shown to retain its ability to bind glutathione affinity matrix<sup>9</sup> and therefore should still be eluted efficiently in the presence of excess glutathione, which was not the case for the protein aggregates we observed (Fig. 5B). Furthermore, when recombinant GST was expressed on its own in BL21(DE3) cells and subjected to Factor Xa proteolysis under conditions identical to those used for digestion of GST-nsp1, a large fraction of GST could still be eluted from the affinity resin with glutathione (Fig. 5B). The nsp1 moiety is therefore the most likely cause of the observed aggregation, which was not related to the presence of the affinity resin, since it was also detected upon Factor Xa digestion of GST-nsp1 in solution (data

not shown). Taken together, these results prompted us to better characterize the properties of purified GST-nsp1. The fusion protein was eluted from GS4B and dialyzed against a wide range of buffers with different pH and ionic strength, all of which contained 1 mM CaCl<sub>2</sub> and 50 mM L-Arg (essential components of Factor Xa buffer). The aggregation state of the fusion protein was then assessed by dynamic light scattering, a method for determination of particle size in solution. The hydrodynamic radius of GST-nsp1 particles indicated that the fusion protein is largely present as “soluble aggregates” in the majority of buffers tested, or completely precipitated under certain conditions (data not shown). In view of the limitations on buffer additives compatible with Factor Xa activity outlined above, further optimization aimed at improving GST-nsp1 solubility was not pursued.



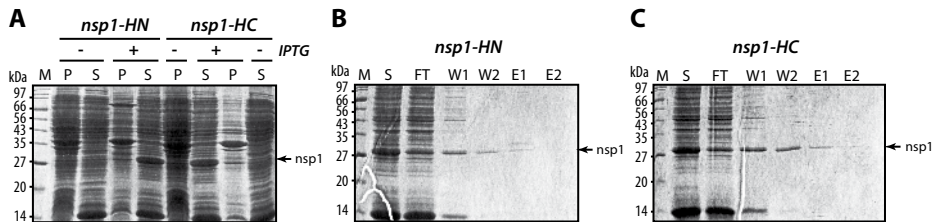
**Figure 5.** Protein precipitation during Factor Xa cleavage of GST-nsp1. (A) Expression and purification of GST-nsp1, followed by on-resin cleavage of the GST tag by Factor Xa Protease were performed essentially as in Fig. 4A, apart from the absence of CHAPS from the cleavage buffer. Cleaved nsp1 was recovered from the buffer fraction after pelleting of GS4B by centrifugation, and unprocessed and/or resin-associated proteins were eluted from GS4B with 1x LSB. Equal volumes of LSB eluate and cleaved (C) protein fractions were analyzed by SDS-PAGE and Coomassie blue staining. (B) GST-nsp1 was expressed in BL21(DE3) or C43(DE3) cells. GST was expressed in BL21(DE3) cells that were transformed with empty pGEX-5X-3 vector. GS4B-bound GST-nsp1 and GST only were subjected to proteolysis with Factor Xa Protease in FXa buffer and cleaved proteins (designated with C) were recovered from the buffer fraction. Proteins that remained bound to GS4B after the proteolytic digestion were eluted once with 10 mM reduced glutathione in 20 mM Tris pH=7.5, 150 mM NaCl (E), followed by elution with 1x LSB. Equal volumes of the C, E and LSB fractions were analyzed by SDS-PAGE followed by Coomassie Blue staining.

### Expression and purification of hexahistidine-tagged nsp1

The difficulties encountered during expression and purification of recombinant GST-nsp1 were seemingly related to the inherent instability and/or improper folding of nsp1 when fused to the C-terminus of GST. Taking this into consideration, fusion to another large protein tag with solubility-enhancing properties was not tested. Instead, a (His)<sub>6</sub> tag was fused to the N- or C-terminus of nsp1. A (His)<sub>6</sub> tag allows for convenient one-step protein purification by immobilized metal affinity chromatography. In general, this small tag has little effect on protein folding or stability, and consequently, it rarely needs to be removed from the target protein after purification. However, it does not alter the solubility properties of the target protein. It was therefore surprising that, upon IPTG induction



of (His)<sub>6</sub>-tagged nsp1 expression in BL21(DE3) cells transformed with pDEST14-nsp1-HN or pDEST14-nsp1-HC, robust expression of a 30-kDa protein was detected only in the induced cultures, and was present predominantly in the soluble fractions of bacterial lysates (Fig. 6A). Western blotting with an nsp1-specific antibody confirmed the identity of the recombinant proteins produced in induced cultures (data not shown). Both nsp1-HN and nsp1-HC bound very poorly to Ni-NTA agarose, however, and the majority of both recombinant proteins were found in the flow-through fractions (Fig. 6, B and C). Similar results were obtained when a different metal affinity matrix (HIS-Select HF Nickel Affinity Gel) was used, and washing out unbound proteins in the absence of imidazole did not improve the recovery of nsp1-HN or nsp1-HC neither, but increased the amount of contaminants in the elution fractions (data not shown). The PCP $\beta$  domain of nsp1 serves to autoproteolytically release the C-terminus of the protein from the rest of pp1a during EAV infection, cleaving a Gly-Gly sequence<sup>14</sup>, and since the protease domain is active in *E. coli*, it might have cleaved off the C-terminal (His)<sub>6</sub> tag in the nsp1-HC protein. Expression of an active-site protease mutant (C164S) from pDEST12-nsp1-HC, however, did not improve the binding of nsp1-HC to metal affinity matrices (data not shown). These results suggest that neither the N-terminal nor the C-terminal (His)<sub>6</sub> tag is exposed for binding metal ligands under the experimental conditions used, while a control (His)<sub>6</sub>-tagged protein could be purified efficiently under identical conditions (data not shown). Nevertheless, both nsp1-HN and nsp1-HC seemed to be stably folded in bacterial cells in the absence of solubility-enhancing fusion tags indicating to importance of authentic terminal sequences for folding. The protein's N-terminal subdomain is a ZF (Fig. 1B), with the first of predicted zinc-coordinating residues located relatively close (24 amino acids)



**Figure 6.** Expression and purification of His-tagged nsp1. (A) BL21(DE3) cells transformed with pDEST14-nsp11-HN or pDEST14-nsp11-HC were cultured until  $OD_{600}$  of 0.7, recombinant protein expression was induced by the addition of 0.5 mM IPTG, and the cultures were incubated at 30°C. Lysates were obtained from non-induced or IPTG-induced cells harvested 18 hours postinduction. After centrifugation of total cell lysates, the insoluble pellet (P) fractions were resuspended in 1x LSB, and equal volumes of pellet and supernatant (S) fractions were separated on a 12.5 % polyacrylamide gel by SDS-PAGE. Proteins were visualized by subsequent staining of the gel with Coomassie blue. (B, C) The soluble fraction of the total bacterial lysates from induced cells in (A) was incubated with Ni-NTA Agarose. The beads were pelleted by centrifugation, and the flow-through (FT) was collected for analysis of unbound proteins. Contaminants were then removed by washing the beads twice with buffer B containing 10 mM imidazole. Bound nsp1-HN (B) or nsp11-HC (C) was subsequently eluted with buffer B supplemented with 250 mM imidazole in two steps. Proteins from the pellet and soluble fractions of the total bacterial lysates, as well as the FT, wash (W) and elution (E) steps were analyzed by SDS-PAGE and Coomassie Blue staining.

to nsp1's N-terminus which might interact with other residues for folding to proceed properly. Also, in view of the lack of *in trans* PCP $\beta$  proteolytic activity, it has been proposed that after co-translational self-processing, the C-terminal residues of nsp1 might fold back into the protease active site, thereby blocking subsequent cleavage events<sup>24</sup>. Another viral autoprotease, the alphavirus Sindbis virus core protein, is rendered proteolytically inactive following autocatalytic *cis* cleavage in this manner<sup>5</sup>. A C-terminal (His)<sub>6</sub> tag would thus also be inaccessible for binding to affinity resins. To overcome these intrinsic problems with tag accessibility, attempts were made to purify individual (His)<sub>6</sub>-tagged nsp1 subdomains of nsp1. Constructs encoding only the ZF and PCP $\alpha$  domains (amino-acids 1 to 140 or 1 to 156 from EAV pp1a) fused to a C-terminal (His)<sub>6</sub> tag expressed very poorly in bacterial cells and were mostly insoluble, while expression of an N-terminally (His)<sub>6</sub>-tagged PCP $\beta$  (amino-acids 151 to 260 from EAV pp1a) was not detected at all in induced cultures (data not shown).

### **Importance of nsp1 autoproteolytic processing for recombinant protein stability**

In the course of this project, the crystal structure of nsp1 $\alpha$  from another arterivirus, porcine reproductive and respiratory syndrome virus (PRRSV; strain XH-GD) was reported by Sun and co-workers<sup>16</sup>. The N-terminal part of PRRSV pp1a/pp1ab also contains a ZF domain followed by two PCP domains. Unlike its ortholog in EAV, PCP $\alpha$  is proteolytically active, mediating the self-release of nsp1 $\alpha$  from the viral replicase polyproteins, while the PCP $\beta$  domain cleaves nsp1 $\beta$  off<sup>6</sup>. Homology between EAV nsp1 and PRRSV nsp1 $\alpha$  is thus limited to the first two subdomains of EAV nsp1. Nevertheless, PRRSV nsp1 $\alpha$  is also essential for sg mRNA production in virus-infected cells<sup>11</sup>, suggesting similar functions of the two distantly related proteins in the replicative cycle of arteriviruses. In the report by Sun and co-authors, PRRSV nsp1 $\alpha$  was purified after expression of N-terminally (His)<sub>6</sub>-tagged nsp1 $\alpha$ -nsp1 $\beta$  sequence in *E. coli* in order to obtain nsp1 $\alpha$  in its self-processed form. Interestingly, a number of nsp1 $\alpha$  C-terminal residues were indeed found to reside in the substrate-binding pocket of the protease domain in a well-stabilized form that would preclude further proteolytic reactions. Thus, autoproteolytic processing of PRRSV nsp1 $\alpha$  might very well be a prerequisite for correct folding and stability of the protein – a notion further supported by the insolubility of a protease active-site mutant, C76S, when expressed in *E. coli* under a variety of conditions.

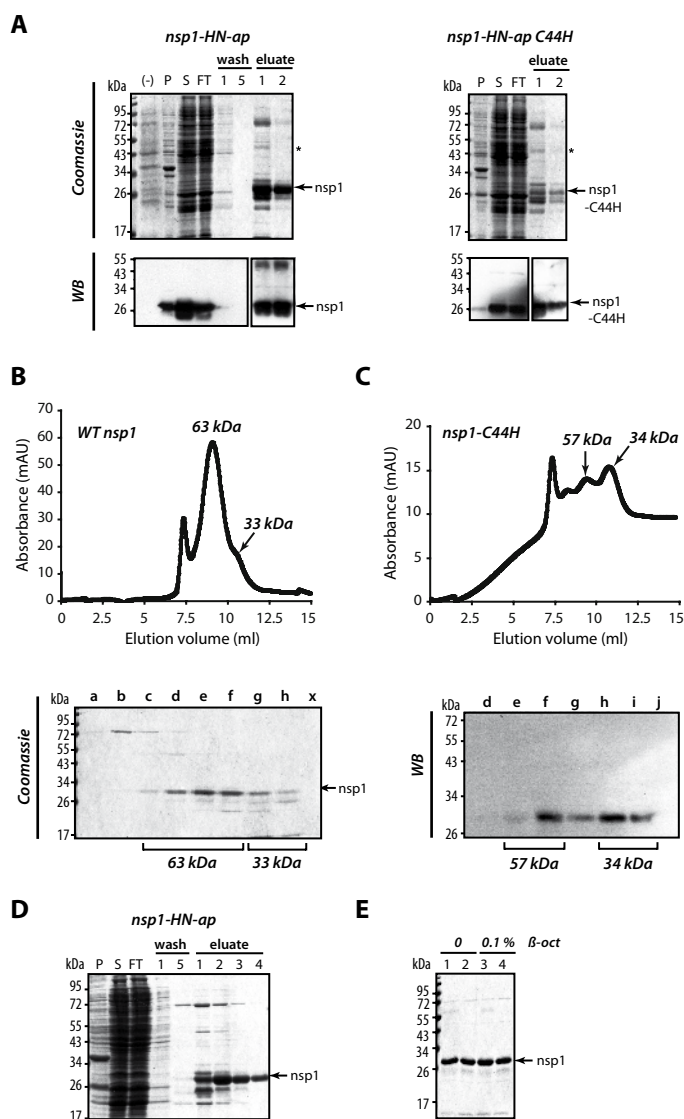
Although the problems we encountered while attempting to purify (His)<sub>6</sub>-tagged EAV nsp1 did not seem related to the protein's solubility but rather to tag accessibility, we could not exclude that nsp1-HN and nsp1-HC were also present as soluble aggregates in lysates fractions, which could also account for poor binding to metal affinity resins. We therefore set out to determine whether bacterially expressed self-cleaved nsp1 could be purified with more success. To this end, the N-terminal 429 amino-acids from EAV pp1a, fused to an N-terminal (His)<sub>6</sub> tag (see Fig. 1C), were cloned in pDEST14, resulting in expression plasmid pDEST14-nsp1-2N-HN. Nsp1 expressed from similar constructs was previously shown to efficiently mediate its autoproteolysis both *in vitro* and *in vivo*<sup>14,18</sup>,

and the difference in molecular mass between the precursor protein (47 kDa) and native nsp1 (30 kDa) is sufficiently large to be easily seen by SDS-PAGE. Recombinant (His)<sub>6</sub>-tagged nsp1 obtained from this construct is referred to as nsp1-HN-ap to distinguish it from nsp1-HN, although the primary structure of the two proteins is identical.

The wild-type pDEST14-nsp1-2N-HN and a mutant derivative carrying a C44H substitution in nsp1 (a mutation of a zinc-coordinating residue in nsp1 that renders sg mRNA production undetectable when introduced into the infectious cDNA clone of EAV) were transformed in C41(DE3) cells, and protein expression and purification were performed following a protocol similar to that used by Sun and co-workers for the purification of PRRSV nsp1 $\alpha$ . Briefly, transformed cells were grown at 37°C to mid-log phase and the culture temperature was lowered to 16°C during protein production, which was induced by addition of 1 mM IPTG. The levels of recombinant protein in the induced cultures were not sufficiently high to be easily detected by Coomassie staining (Fig. 7A, top panels), but Western blotting with an nsp1-specific antibody revealed robust expression of both nsp1-HN-ap and its mutant derivative, nsp1-HN-ap C44H (Fig. 7A, bottom panels). Both proteins were found predominantly in the supernatant fractions of bacterial lysates and exclusively in self-processed form, since virtually no precursor proteins were detected. These results suggest that nsp1 autoproteolysis with this combination of expression construct and expression conditions is very efficient in *E. coli*, despite the low temperature of protein production. The C44H mutation also does not seem to influence the efficiency of nsp1 self-processing, in line with previous results showing that mutations of zinc-coordinating residues do not affect nsp1 autoprocessing *in vitro* (Chapter 5 of this thesis).

Wild-type nsp1-HN-ap could be efficiently purified by metal affinity chromatography with Talon resin (charged with cobalt ions), albeit in the presence of a number of abundant protein contaminants (Fig. 6A, top left panel). Also, some of the protein was still found in the fraction of proteins that did not bind to the affinity resin. Nevertheless, this protocol resulted in the purification of ~0.2 mg recombinant nsp1-HN-ap from 1 l of bacterial culture (data not shown). The yield of nsp1-HN-ap C44H, however, was much lower (Fig. 7A, top right panel), and immunoblot analysis of fractions obtained during purification revealed that most of the mutant recombinant protein did not bind the Talon affinity resin (Fig. 7A, bottom right panel). The C44H substitution might interfere with zinc binding and thus adversely affect the conformational stability of the N-terminal part of nsp1, and possibly of the whole protein, if interdomain interactions are required for its correct folding. This might not impede the PCP $\beta$  domain-mediated autoproteolytic release of nsp1, which most probably occurs co-translationally<sup>14,15</sup>, but could result in protein aggregation that could, in turn, render the (His)<sub>6</sub> inaccessible for metal binding.

In order to improve sample purity, gel-filtration chromatography of pooled eluates containing recombinant nsp1-HN-ap (Fig. 6A) was performed in 20 mM MES pH=7.0, 500 mM NaCl, 20 % glycerol, 1 mM DTT. The bulk of wild-type nsp1-HN-ap eluted in a single peak with an estimated molecular mass of 63 kDa (Fig. 7B), consistent with a dimer. Column fractions from the single elution peak preceding that of dimeric nsp1 contained largely a ~75-kDa contaminant (Fig. 7B, lower panel, a-b), and higher oligomeric forms



**Figure 7.** Expression and purification of autoproteolytically processed His-tagged nsp1. (A) C41(DE3) cells were transformed with wild-type pDEST14-nsp1-2N-HN or a mutant derivative of the plasmid carrying a C44H substitution in nsp1. Induction of recombinant protein expression and purification of His-tagged proteins was performed as described in Materials and Methods. Samples obtained from total lysate of non-induced [(-)] cells, pellet (P) and supernatant (S) fractions of total lysates from IPTG-induced cells, and fractions obtained during purification (FT, flow-through) were resolved by SDS-PAGE. The gels were stained with Coomassie blue (top panels) or blotted onto Hybond-P PVDF membrane (GE Healthcare) by semi-dry Western blotting (WB, bottom panels). The membranes were subsequently incubated with an nsp1-specific monoclonal antibody (12A4), followed by an HRP-conjugated secondary antibody (DAKO) and an ECL-Plus kit (GE Healthcare) for detection. The expected position of the unprocessed nsp1-2N-HN protein (predicted molecular mass: 46.6 kDa) is indicated with an asterisk. (B, C) Top: Gel-filtration profile of wild-type nsp1-HN-ap (B) or nsp1-HN-ap C44H (C) obtained upon size exclusion chromatography of the pooled

eluates in (A) on a Superdex-75 column. Blue dextran (2,000 kDa) was used to determine the column void volume, and the elution volumes of conalbumin (75 kDa), ovalbumin (43 kDa) and chymotrypsinogen A (25 kDa) filtered through Superdex-75 in the same buffer conditions were fitted to a linear equation ( $R^2=0.97$ ). The standard curve was then used to estimate the molecular masses of nsp1 elution peaks that were within the linear range. Bottom: peak fractions from top panels were analyzed by SDS-PAGE and Coomassie Blue staining (B) or Western blotting with the nsp1-specific antibody (C). (D) Wild-type nsp1-HN-ap was purified as in (A), with the following modifications: addition of 5 mM  $\beta$ -mercaptoethanol to buffer C during bacterial lysis and Talon purification, and a single-step elution with 100 mM imidazole, followed by three elution steps with 250 mM imidazole in buffer C. Fractions from cell lysis and Talon purification analyzed by SDS-PAGE and Coomassie blue staining are denoted as in (A). (E) Eluates 2 – 4 from (D) were pooled and subjected to size-exclusion chromatography on a Superdex-75 column in 20 mM MES pH=7.0, 1M NaCl, mM DTT. Nsp1-containing peak fractions were concentrated in the same buffer, or in the presence of 0.1 %  $\beta$ -octylglucoside ( $\beta$ -oct). Equal sample volumes obtained immediately after concentration (1, 3) or following overnight incubation of the concentrated samples at 4°C (2, 4) were resolved by SDS-PAGE and visualized by Coomassie blue staining.

of nsp1 were thus unlikely to be present in solution. The resolution of the gel filtration column was, however, insufficient to achieve a substantially improved purity of recombinant nsp1. Despite the low protein yield and purity, the fractions eluted from the Talon resin after purification of nsp1-HN-ap C44H (Fig. 7A, left panel) were also analyzed by gel filtration, to assess whether the C44H mutation is associated with a change in oligomeric state. The elution profile of nsp1-HN-ap C44H was much less symmetric due to the complexity of the sample, but the mutant protein was only detected in fractions from two elution peaks with approximately equal areas and estimated molecular masses of 57 kDa and 34 kDa (Fig. 7C). The C44H thus seems to destabilize, but not completely prevent nsp1 self-association. It is difficult to establish, however, whether this may be attributable to an importance of the ZF domain for dimerization, or to a more general defect in the overall fold of nsp1 (see above). Nevertheless, recombinant nsp1 purified by two different protocols forms dimers in solution (Fig. 4B and 7B), even in the presence of 1 mM DTT and high salt concentrations (500 mM NaCl, Fig. 7B). It is therefore unlikely that the observed dimerization is caused by tag sequences in the recombinant protein. Also, in view of nsp1's homo-oligomerization in infected cells<sup>19</sup>, the ability of recombinant nsp1 to form dimers might be of functional significance for its role(s) in the EAV replicative cycle.

Though nsp1-HN-ap was efficiently resolved by size-exclusion chromatography in 20 mM MES pH=7.0, 500 mM NaCl, 20 % glycerol, 1 mM DTT, the protein was unstable in this buffer, since precipitation was observed during concentration of nsp1-containing chromatography fractions. The loss of protein was even more pronounced after overnight storage at 4°C, or when the sample buffer was exchanged with one of a lower ionic strength (20 mM MES pH 7.0, 200 mM NaCl, 50% glycerol). In an attempt to obtain a purer and more stable preparation of recombinant nsp1-HN-ap, the protein was expressed and purified as above, with minor adjustments to the original protocol (described in detail in Materials and Methods). Beta-mercaptoethanol was added at 5 mM to buffer C prior to bacterial lysis and Talon affinity purification to prevent the formation of disulfide bonds that might contribute to the high level of protein contaminants co-purifying with nsp1. Also, elution

of proteins bound to the Talon resin was performed as follows: first, proteins that were bound with low affinity were eluted with 100 mM imidazole in buffer C, followed by three elution steps with 250 mM imidazole in buffer C. Most of the contaminants were eluted in the 100 mM imidazole step, together with a relatively low proportion of recombinant nsp1-HN-ap. The nsp1-containing eluate fractions obtained in this way were of higher purity (Fig. 7D). The pooled eluates were loaded onto a Superdex-75 column, gel-filtration chromatography was carried out in 20 mM MES pH=7.0, 1M NaCl, 20 % glycerol, 5 mM DTT. The peak column fractions that contained nsp1 were split in two samples that were subsequently concentrated at 4°C by ultrafiltration in the absence of additives, or in the presence of 0.1 %  $\beta$ -octylglucoside. There was no detectable protein precipitation during ultrafiltration in both buffers, and SDS-PAGE showed similar levels of recombinant nsp1 present in the two samples immediately after protein concentration, as well as following overnight incubation at 4°C (Fig. 7E). The yield of nsp1-HN-ap after Talon affinity purification and after gel filtration and was estimated at 0.3 mg and 0.25 mg/ l culture, with the minor loss of protein likely due to exclusion of gel-filtration fractions that contained a low level of nsp1 and high levels of contaminating proteins (data not shown). The inclusion of an additional purification step, i.e. ion exchange chromatography, might remove co-purifying proteins more efficiently to obtain a sample that is sufficiently pure for structure determination.

## CONCLUDING REMARKS

Despite recent insights into successful strategies for the expression and purification of recombinant proteins emerging from large-scale structural genomics studies (see e.g. reference<sup>7</sup>), optimal conditions for each protein still have to be determined empirically. Our experience greatly underscores the importance of the choice of fusion tag that can be facilitated by at least some prior knowledge about the domain structure and/or biochemical properties of the target protein. The extensive aggregation we observed upon expression of nsp1 as a C-terminal GST fusion might be due to a number of factors, such as e.g. improper folding of the ZF domain when the N-terminus of nsp1 was blocked by its fusion to GST, or the formation of GST dimers forcing nsp1 dimerization before the protein has reached its stably folded state. Ultimately, successful purification of soluble recombinant nsp1 was achieved with nsp1 which was autoproteolytically released from a precursor polypeptide in *E. coli*, an observation that is in line with results of other studies in which soluble viral proteins that are derived from proteolytic processing of longer polypeptides were obtained<sup>5,16</sup>. It is thus tempting to suggest that mimicking the natural way of protein production may be a generally applicable approach for obtaining soluble recombinant viral proteins. This strategy could also constitute an important addition to the widely-used approaches of expressing subdomains and/or truncated proteins for structural and functional studies.

## ACKNOWLEDGEMENTS

We thank Martijn van Hemert and Sjoerd van den Worm for helpful discussions, and Linda Struijk and Els van der Meijden for the pGEX-5X-3 plasmid.

## REFERENCES

1. **Abeliovich, H. and J. Shlomai.** 1995. Reversible oxidative aggregation obstructs specific proteolytic cleavage of glutathione S-transferase fusion proteins. *Anal.Biochem.* **228**:351-354
2. **Arakawa, T. and K. Tsumoto.** 2003. The effects of arginine on refolding of aggregated proteins: not facilitate refolding, but suppress aggregation. *Biochem.Biophys.Res.Commun.* **304**:148-152
3. **Beerens, N., B. Selisko, S. Ricagno, I. Imbert, L. van der Zanden, E. J. Snijder, and B. Canard.** 2007. De novo initiation of RNA synthesis by the arterivirus RNA-dependent RNA polymerase. *J.Virol.* **81**:8384-8395
4. **Blackwell, J. R. and R. Horgan.** 1991. A novel strategy for production of a highly expressed recombinant protein in an active form. *FEBS Lett.* **295**:10-12
5. **Choi, H. K., L. Tong, W. Minor, P. Dumas, U. Boege, M. G. Rossmann, and G. Wengler.** 1991. Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion. *Nature* **354**:37-43
6. **den Boon, J. A., K. S. Faaberg, J. J. Meulenberg, A. L. Wassenaar, P. G. Plagemann, A. E. Gorbalenya, and E. J. Snijder.** 1995. Processing and evolution of the N-terminal region of the arterivirus replicase ORF1a protein: identification of two papainlike cysteine proteases. *J.Virol.* **69**:4500-4505
7. **Graslund, S., P. Nordlund, J. Weigelt, B. M. Hallberg, J. Bray, O. Gileadi, S. Knapp, U. Oppermann, C. Arrowsmith, R. Hui, J. Ming, S. dhe-Paganon, H. W. Park, A. Savchenko, A. Yee, A. Edwards, A. Vincentelli, C. Cambillau, R. Kim, S. H. Kim, Z. Rao, Y. Shi, T. C. Terwilliger, C. Y. Kim, L. W. Hung, G. S. Waldo, Y. Peleg, S. Albeck, T. Unger, O. Dim, J. Priluski, J. L. Sussman, R. C. Stevens, S. A. Lesley, I. A. Wilson, A. Joachimiak, F. Collart, I. Dementieva, M. I. Donnelly, W. H. Eschenfeld, Y. Kim, L. Stols, R. Wu, M. Zhou, S. K. Burley, J. S. Emtage, J. M. Sauder, D. hompson, K. ain, J. Luz, T. Gheyi, F. Zhang, S. Atwel, S. C. Almo, J. B. Bonanno, A. Fiser, S. Swaminathan, F. W. Studier, M. R. Chance, A. Sali, T. B. Acton, R. Xiao, L. Zhao, L. C. Ma, J. F. Hunt, L. Tong, K. Cunningham, M. Inouye, S. Anderson, H. Janjua, R. Shastry, C. K. Ho, D. Wang, H. Wang, M. Jiang, G. T. Montelione, D. I. Stuart, R. J. wens, S. aenke, A. Schutz, U. Heinemann, S. Yokoyama, K. Bussow, and K. C. Gunsalus.** 2008. Protein production and purification. *Nat.Meth.* **5**:135-146
8. **Ignatova, Z. and L. M. Gierasch.** 2006. Inhibition of protein aggregation in vitro and in vivo by a natural osmoprotectant. *Proc.Natl.Acad.Sci.U.S.A* **103**:13357-13361
9. **Kaplan, W., P. Husler, H. Klump, J. Erhardt, N. Sluis-Cremer, and H. Dirr.** 1997. Conformational stability of pGEX-expressed *Schistosoma japonicum* glutathione S-transferase: a detoxification enzyme and fusion-protein affinity tag. *Protein Sci.* **6**:399-406
10. **Kataeva, I., J. Chang, H. Xu, C. H. Luan, J. Zhou, V. N. Uversky, D. Lin, P. Horanyi, Z. J. Liu, L. G. Ljungdahl, J. Rose, M. Luo, and B. C. Wang.** 2005. Improving solubility of *Shewanella oneidensis* MR-1 and *Clostridium thermocellum* JW-20 proteins expressed into *Escherichia coli*. *J.Proteome. Res.* **4**:1942-1951
11. **Kroese, M. V., J. C. Zevenhoven-Dobbe, J. N. Bos-de Ruijter, B. P. Peeters, J. J. Meulenberg, L. A. Cornelissen, and E. J. Snijder.** 2008. The nsp1alpha and nsp1 papain-like autoproteases are essential for porcine reproductive and respiratory syndrome virus RNA synthesis. *J.Gen.Virol.* **89**:494-499
12. **Miroux, B. and J. E. Walker.** 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J.Mol.Biol.* **260**:289-298



13. **Sawicki, S. G. and D. L. Sawicki.** 1995. Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. *Adv.Exp.Med.Biol.* **380**:499-506
14. **Snijder, E. J., A. L. Wassenaar, and W. J. Spaan.** 1992. The 5' end of the equine arteritis virus replicase gene encodes a papainlike cysteine protease. *J.Virol.* **66**:7040-7048
15. **Snijder, E. J., A. L. Wassenaar, and W. J. Spaan.** 1994. Proteolytic processing of the replicase ORF1a protein of equine arteritis virus. *J.Virol.* **68**:5755-5764
16. **Sun, Y., F. Xue, Y. Guo, M. Ma, N. Hao, X. C. Zhang, Z. Lou, X. Li, and Z. Rao.** 2009. Crystal structure of porcine reproductive and respiratory syndrome virus leader protease nsp1alpha. *J.Virol.* **83**:10931-10940
17. **Thomas, J. G. and F. Baneyx.** 1996. Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpressing Heat-shock proteins. *J.Biol.Chem.* **271**:11141-11147
18. **Tijms, M. A., D. D. Nedialkova, J. C. Zevenhoven-Dobbe, A. E. Gorbalenya, and E. J. Snijder.** 2007. Arterivirus subgenomic mRNA synthesis and virion biogenesis depend on the multifunctional nsp1 autoprotease. *J.Virol.* **81**:10496-10505
19. **Tijms, M. A. and E. J. Snijder.** 2003. Equine arteritis virus non-structural protein 1, an essential factor for viral subgenomic mRNA synthesis, interacts with the cellular transcription co-factor p100. *J.Gen.Virol.* **84**:2317-2322
20. **Tijms, M. A., L. C. van Dinten, A. E. Gorbalenya, and E. J. Snijder.** 2001. A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus. *Proc.Natl.Acad.Sci.U.S.A* **98**:1889-1894
21. **van den Born, E., A. P. Gultyaev, and E. J. Snijder.** 2004. Secondary structure and function of the 5'-proximal region of the equine arteritis virus RNA genome. *RNA* **10**:424-437
22. **van Hemert, M. J., A. H. de Wilde, A. E. Gorbalenya, and E. J. Snijder.** 2008. The in vitro RNA synthesizing activity of the isolated arterivirus replication/transcription complex is dependent on a host factor. *J.Biol.Chem.* **283**:16525-16536
23. **Vera, A., N. Gonzalez-Montalban, A. Aris, and A. Villaverde.** 2007. The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnol.Bioeng.* **96**:1101-1106
24. **Ziebuhr, J., E. J. Snijder, and A. E. Gorbalenya.** 2000. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J.Gen.Virol.* **81**:853-879



# Chapter 8

**General discussion |**




## DISSECTING THE ROLES OF REPLICASE SUBUNITS IN NIDOVIRUS RNA SYNTHESIS

Plus-strand RNA virus replication requires cytoplasmic RNA-templated RNA synthesis - a process foreign to cells and thus accomplished by a virus-encoded RdRp, which is produced upon translation of the viral genome following its uncoating in the infected cell. As is clear from the *in vitro* activity assays developed for a variety of recombinant viral RdRps<sup>3,4,17,54,61,81,85,103</sup>, this enzymatic activity can be sufficient for catalysis of RNA polymerization. Apart from this critical process, however, +RNA viruses also need to ensure the specificity of viral RNA amplification and encapsidation in an environment where cellular mRNAs are abundant, as well as to evade cellular defense mechanisms. These requirements have necessitated the evolution of complex regulatory circuits that control +RNA replication. They are commonly mediated by virus-encoded domains and can also make use of cellular proteins. The viral regulators are typically produced as separate protein units but may also be part of multidomain proteins. In view of their (usual) absence from virions, these proteins are called “nonstructural”, differentiating them from the structural components of virus particles. Collectively, they are commonly referred to as “the viral replicase”. The coordination of various processes in +RNA replication is directed by replicase subunits, via their enzymatic activities and/or non-catalytic modulatory functions. The replicase proteins of RNA viruses are often multifunctional, a property likely stemming from the limited coding capacity of RNA virus genomes, which has in turn been attributed to the low fidelity of their replication machinery<sup>15,27</sup>.

Nidoviruses arguably encode the most complex replicase among RNA viruses known to date<sup>19</sup>, and between 13 and 16 individual viral nonstructural proteins (nsps) are produced in nidovirus-infected cells<sup>90,96,97,104,107,108</sup>. A replicative machinery of such complexity may have been necessitated by the nidovirus gene expression strategy, which includes the synthesis of an extensive set of subgenomic (sg) mRNAs, and by the need to maintain the exceptionally large genomes characteristic of most *Nidovirales* representatives<sup>19</sup>. Most nsps localize to virus-induced networks of modified membranes found in the perinuclear region of nidovirus-infected cells, which are thought to be the sites of viral RNA synthesis<sup>32,53,56,77</sup>. A large fraction of nidovirus nsps have predicted and/or demonstrated enzymatic activities. Most of these have distant protein homologs in the cellular world, and a smaller subset also among RNA viruses<sup>9,18,24,26,42,74</sup>. Understanding the roles of individual replicase subunits in the nidovirus replicative cycle will undoubtedly shed light on how are nidovirus RNA-synthesizing complexes assembled and what are the determinants of their diverse activities. This knowledge, in turn, will be instrumental in designing strategies to combat known and emergent diseases caused by nidoviruses.

More than a decade ago, a reverse genetics system was developed for equine arteritis virus (EAV), the arterivirus prototype and the nidovirus with the smallest genome thus far<sup>95</sup>. This system has greatly benefitted research on the fundamental aspects of nidovirus molecular biology. Important insights into the unique process employed by arteri- and coronaviruses to generate a nested set of 5'- and 3'-coterminally sg mRNAs have been





obtained by dissecting the roles of EAV regulatory RNA sequences, like TRSs and leader TRS hairpin<sup>48,51,52,91,92,100</sup>. Having the most “compact” nidovirus replicase, EAV is commonly considered as a powerful model system for functional, structural and genetic studies of the elaborate nidovirus replicative apparatus. Among the most significant findings obtained with the EAV reverse genetics system was the discovery that nidovirus genome replication can be uncoupled from sg mRNA production via mutations in replicase subunits<sup>86,95,101</sup>. This observation strongly suggests that these two processes are carried out by RNA-synthesizing complexes of (partially) different protein composition.

The studies described in this thesis focused on the role of nidovirus replicase subunits in coupling different processes in the viral replicative cycle. Two very different proteins were examined: one encoded in ORF1a and specific for arteriviruses (nsp1), and another one encoded in ORF1b and containing a domain that is conserved across a broad spectrum of nidoviruses (NendoU). An experimental approach combining mutagenesis of replicase subunits (using full-length cDNA clones) with appropriate *in vitro* biochemical assays that address specific functions of individual nsps can yield biologically relevant insights into the functions of nidovirus nonstructural proteins. This approach was employed in the research outlined in Chapters 3 and 4 to gain more insight into the role of the conserved NendoU domain in the EAV replicative cycle and to examine, for the first time, the *in vitro* endoribonuclease activity of the arterivirus NendoU-containing replicase subunit. In Chapters 5 and 6, we explored the regulatory roles of EAV nsp1, so far the sole nidovirus protein with a demonstrated specific role in sg mRNA production. Our data, obtained using reverse and forward genetics, revealed that nsp1 is a key coordinator of genome replication, sg mRNA synthesis, and virus production. To conduct research that could help unravel the molecular mechanisms underlying nsp1 function, we have developed a protocol for purification of recombinant nsp1 from *E. coli* (Chapter 7). This accomplishment occurred at the final stages of the studies presented in this thesis, leaving it to others to fully exploit its potential. In the following sections, our data on NendoU and nsp1 are discussed in substantial detail in the framework of relevant findings obtained by others, and potential future research directions are outlined.

## THE NIDOVIRUS REPLICATIVE ENDORIBONUCLEASE

Endoribonucleases are ubiquitous in both prokaryotic and eukaryotic cells and play important roles in a diverse set of processes related to the maturation of tRNAs and mRNAs, as well as the regulated degradation of cellular transcripts<sup>37,44,80</sup>. Both DNA and RNA viruses employ virus-encoded endoribonucleases to perform various functions, such as inhibition of cellular mRNA translation or post-transcriptional regulation of viral gene expression<sup>82,83,89</sup>. Among these is also the sole endoribonuclease other than NendoU encoded by +RNA viruses – the pestivirus E(rns), an essential structural component of pestivirus particles that also acts as an interferon antagonist<sup>23,39,40,70</sup>.

The NendoU domain is conserved among the different nidovirus groups both at the amino acid sequence level, as well as in terms of its relative position in the C-terminal part of the pp1ab replicase polyprotein. NendoU was postulated to be involved in a nidovirus-specific step of the viral replicative cycle, based on its genetic segregation with the core RNA-synthesizing enzymes and the lack of NendoU counterparts in the replicase genes of other known RNA virus groups<sup>12,74</sup>. Also, the replicase genes of nidoviruses with large genomes harbor additional RNA-processing domains whose functions may be interconnected with that of NendoU<sup>19,74</sup>. Structural and functional studies on coronavirus NendoU<sup>5-7,22,26,30,60,105</sup> had initially verified its proposed endoribonuclease activity and established several mechanistic aspects of coronavirus NendoU-mediated RNA cleavage *in vitro*. By contrast, the biochemical characterization of NendoU from other nidovirus groups lagged behind. We were able to overcome the considerable technical difficulties that were associated with obtaining a recombinant form of nsp11, the NendoU-containing replicase subunit from EAV. The availability of recombinant nsp11 allowed us to initiate the biochemical characterization of the arterivirus NendoU, which was carried out in comparison with a coronavirus ortholog (Chapter 4). This *in vitro* functional analysis was complemented by a reverse genetic study of the EAV NendoU domain, the results of which underscored the importance of this enzyme for arterivirus replication (Chapter 3).

### Substrate specificity and regulation of NendoU activity

Recombinant forms of NendoU-containing arteri- and coronavirus replicase subunits exhibit broad substrate specificity *in vitro*, processing both single-stranded and double-stranded RNA substrates 3' of pyrimidines. The modest preference for cleavage at single-stranded uridylylates, noted for both arterivirus nsp11 and coronavirus nsp15 by us and others, seems to be more pronounced for the coronavirus enzyme. This difference is presumably due to amino acid determinants in the C-terminal region of the coronavirus NendoU domain that are seemingly not conserved in the corresponding region of the arterivirus NendoU<sup>(5,6,26,105)</sup> and Chapter 4 of this thesis). Whether these enzymes have similar substrate specificity in the context of viral infection, however, remains to be established.

NendoU enzymes share certain features with bovine pancreatic RNase A, such as the identity of active site residues and mechanism of RNA hydrolysis. Non-catalytic residues of RNase A have also been proposed to contribute to RNA cleavage, and the type of nucleoside immediately downstream of the cleavage site seems to exert significant influence on the rate of substrate hydrolysis, with purines being preferred over pyrimidines<sup>45,58</sup>. A systematic investigation of the importance of nucleotides flanking the cleavage site for RNA hydrolysis by NendoU has not been performed to date, and might provide clues about the natural substrates of this enzyme in infected cells.

Hexamerization of coronavirus nsp15 via intermolecular interactions mediated by its N-terminal domain acts to stabilize the NendoU catalytic site in its active conformation and promotes substrate binding<sup>6,29</sup>. The (local) concentration of nsp15 during coronavi-

rus infection was therefore postulated to act as an allosteric switch in coronavirus NendoU<sup>29</sup>. Arterivirus nsp11, on the other hand, does not seem to form oligomers in solution (Chapter 4 and data not shown) and might thus be similar to its distant cellular homolog XendoU, which is active as a monomer<sup>59</sup>.

Nidovirus nsps containing the NendoU domain mediate efficient RNA processing *in vitro* in the absence of other viral or cellular proteins (<sup>5,26</sup> and Chapter 4). Nevertheless, NendoU activity or substrate specificity may be regulated via protein-protein interactions during infection, and a number of examples of such regulatory interactions have been described for endoribonucleases of both viral and cellular origins. The endoribonuclease activity of bacteriophage T4 RegB is enhanced by the *E. coli* ribosomal protein S1, and the presence of S1 also alters RegB cleavage site specificity. This interaction has been postulated to play a role in ensuring the targeted degradation of early viral transcripts by RegB<sup>14,35,63</sup>. Similarly, the virion host shutoff endoribonuclease of herpes simplex virus 1 associates with polyribosomes to selectively degrade mRNAs that are actively translated<sup>84</sup>. An example of a negative protein regulator is RraA from *E. coli*, which binds to RNase E and disrupts its multimerization, thereby inhibiting the endoribonucleic activity of the enzyme, which is involved in mRNA degradation<sup>8,36</sup>.

The potential modulation of NendoU activity by proteins of viral or cellular origins has not been addressed to date. Large-scale studies of pair-wise interactions between SARS-CoV proteins expressed in cells outside the context of infection have demonstrated an nsp15 self-interaction in a yeast-two-hybrid screen<sup>47</sup>. Assays based on pull-down of *in vitro*-translated replicase subunits with GST-tagged viral proteins have also suggested SARS-CoV nsp15 might interact with nsp8 (the putative RNA primase), nsp9 (an RNA-binding protein), the viral RdRp and the protease domain of nsp3<sup>25</sup>. The results of this study should be interpreted with extreme caution, however, because it is difficult to assess whether proteins derived from *in vitro* translation or fused to GST are properly folded, and GST is also known to impose dimerization on its fusion partners, which may also interfere with the native fold of a protein (see Chapter 7).

It is also unclear how NendoU activity is affected by proteolytic processing of the ORF1b-encoded part of nidovirus replicase polyproteins. Analysis of the cleavage kinetics of EAV pp1ab showed processing of the ORF1b-encoded polypeptides is slow and occurs in no apparent sequential order, giving rise to many processing intermediates, among which nsp10-12 and nsp11-12<sup>97</sup>. Interestingly, cleavage at the nsp9/nsp10 and nsp11/nsp12 junctions of EAV pp1ab was found to be essential for virus viability, while a mutant with impaired processing of the nsp10/nsp11 cleavage site was viable, though severely crippled<sup>96</sup>. The NendoU domain maps to the C-terminal portion of EAV nsp11, and may thus be inactive when fused to downstream nsp12, but capable of processing RNA as part of an nsp10-11 polypeptide. Obtaining recombinant EAV nsp11 precursor polypeptides and analyzing their RNA processing activities *in vitro* could therefore yield important insights into whether NendoU activity is regulated by polyprotein processing. Our attempts to purify recombinant nsp11-12, following the same protocol used for expression and purification of enzymatically active arterivirus nsp11 (see Chapter 4), were



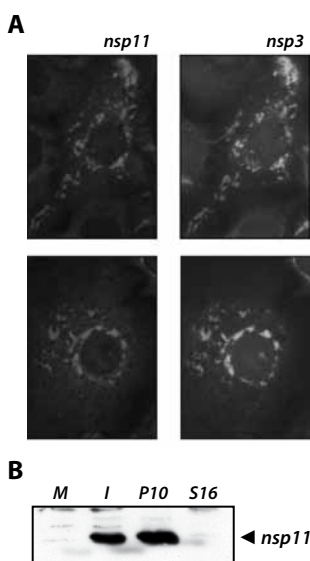
unfortunately unsuccessful, likely due to insolubility or instability of this polypeptide in *E. coli* (data not shown). We have not attempted to obtain recombinant nsp10-11 so far, and this might in fact prove challenging, in view of the high toxicity associated with heterologous expression of both nsp10 and nsp11 (Chapter 4 and unpublished observations of L.C. van Dinten, J.C. Zevenhoven, and E.J. Snijder). Analysis of the *in vitro* duplex unwinding and RNA-processing activities of this precursor, however, may reveal potential mechanisms of regulating not only NendoU, but also the helicase activity of nsp10.

Finally, an alternative approach to assess the importance of nsp11 precursor forms for the replicative cycle of EAV could involve deletion of the nsp11-coding sequence from ORF1b and engineering an EAV cDNA construct that expresses nsp11 from alternative locations in the viral genome, such as ORF1a<sup>93</sup> or a sg mRNA. This experimental setup may also shed light on the significance of the downregulation of expression of NendoU-containing replicase subunits relative to ORF1a-encoded viral nsps.

### **NendoU: a subunit of viral replication complexes?**

Heterologous expression of arterivirus nsp11 and coronavirus nsp15 in both prokaryotic and eukaryotic cells is extremely toxic (<sup>60,105</sup> and Chapter 4). Combined with the lack of toxicity of NendoU mutants defective in RNA processing, this observation suggests that the enzyme may target essential cellular RNA components when it is expressed outside the context of infection. The intracellular localization of NendoU-containing replicase subunits in nidovirus-infected cells may therefore determine their access to RNA substrates. A rabbit polyclonal antibody that recognizes EAV nsp11 (and likely also nsp11-containing precursor polypeptides) was recently raised in our laboratory (D.D. Nedialkova and E.J. Snijder, unpublished data) and used to examine the intracellular distribution of nsp11 in EAV-infected cells. Our preliminary results suggest the protein co-localizes with membrane-anchored replicase subunits at presumed sites of RNA synthesis in perinuclear regions (Fig. 1A), and a similar distribution of SARS-CoV nsp15 has been reported<sup>77</sup>.

Remarkably, arterivirus nsp11 is not predicted to contain hydrophobic or transmembrane domains (data not shown), but the protein efficiently co-sedimented with isolated EAV RNA-synthesizing complexes that are known to be associated with membranes<sup>98</sup> (Fig. 1B). Compartmentalization of NendoU-containing nsps may thus represent a regulatory mechanism for sequestering this enzymatic activity, which is potentially detrimental to cellular mRNAs, into virus-induced membrane compartments. It will be interesting to determine whether nsp11 targeting to the sites of EAV RNA synthesis is mediated by interactions with other viral replicase subunits, and what effect perturbing nsp11 sequestration would have on the viral replicative cycle.



**Figure 1. EAV nsp11: a subunit of viral replication complexes?** (A) Immunofluorescence analysis of nsp11 localization in EAV-infected Vero E6 cells (8 h post-infection). A dual labeling was performed as previously described<sup>94</sup>, using a rabbit polyclonal antibody that recognizes EAV nsp11 (left panels, see text for details) and an Alexa 488-conjugated rabbit antiserum recognizing EAV nsp3 (right panels). (B) Distribution of EAV nsp11 in subcellular fractions from EAV-infected cells. Active viral RNA-synthesizing complexes were isolated from EAV-infected cells by the method of van Hemert *et al.*<sup>98</sup>. Post-nuclear supernatants (PNS) of mock-infected (M) or EAV-infected (I) BHK-21 cells were prepared, and the PNS from EAV-infected cells was fractionated into a 10,000  $\times g$  pellet (P10) and supernatant (S10). The S10 fraction was clarified further by centrifugation at 16,000  $\times g$  to yield S16. The membrane-associated viral RNA-synthesizing complexes are found in the P10 fraction, but require a cytosolic factor from the S16 fraction for *in vitro* activity<sup>98</sup>. Equivalent amounts of M, I, P10 and S16 fractions originating from the same number of cells were separated on SDS-PAGE gels, and the distribution of EAV nsp11 was analyzed by Western blotting with the anti-nsp11 rabbit antiserum.

### Does NendoU have a conserved function in the replicative cycle of arteriviruses and coronaviruses?

A critical role of the NendoU domain-containing nsps in the replicative cycle of both corona- and arteriviruses was established by reverse genetics-based mutagenesis of this domain in EAV (Chapter 3), MHV, and human coronavirus (HCoV) 229E<sup>26,30</sup>. Deletions in the EAV NendoU domain and replacements of conserved Asp residues in NendoU of EAV, MHV and HCoV 229E rendered viral RNA synthesis undetectable when introduced in the corresponding full-length infectious cDNA clones. A more rigorous characterization of these mutants, however, is required to rule out defects in polyprotein processing that could provide an alternative explanation for the observed nonviable phenotypes.

Replacement of any of the three predicted active site residues of NendoU reduced the *in vitro* RNA-processing activities of arterivirus nsp11 and coronavirus nsp15 to levels close to background without affecting protein solubility<sup>(26,30,105)</sup> and Chapter 4). The same substitutions yielded viable, but severely crippled EAV mutants with a specific defect in viral sg mRNA synthesis, accompanied by a dramatic reduction of infectious progeny titers by up to 5 logs (Chapter 3). By contrast, mutagenesis of the MHV NendoU active site resulted in only modest defects in viral RNA synthesis and an up to 1.5 log decrease in progeny titers. Notably, no specific defects in MHV sg mRNA synthesis were observed for these mutants<sup>30</sup>.

Overall, these data underscore the importance of the NendoU domain for efficient viral replication, but also suggest (partially) different roles of NendoU in the replicative cycles of arteri- and coronaviruses. A more thorough characterization of the NendoU mutant phenotypes, such as the quantitative analysis of viral minus-strand RNA species and accumulation levels of viral proteins, may aid in pinpointing steps of the viral replicative

cycle at which NendoU operates. It will also be important to analyze NendoU active-site mutants in relevant *in vivo* infection models, in view of the documented roles of other (putative) nidovirus RNA-processing enzymes in viral pathogenicity<sup>16,62</sup>. Finally, connecting the currently available data from *in vitro* RNA-processing assays and reverse genetics mechanistically requires the identification of the *in vivo* substrates of these enzymes, which will undoubtedly provide clues about the basis for the conservation of this enzymatic activity exclusively in nidoviruses.

### **How can the substrates of NendoU in infected cells be identified?**

Identifying the substrate(s) of NendoU in nidovirus-infected cells is a daunting task and different approaches to tackle it can be envisaged. Isolation of protein complexes containing NendoU either without or in association with its substrates, for example, would open up the possibility of characterizing NendoU RNA-processing *in vitro* and identification of the bound RNA molecules, respectively. Imperative to the success of both approaches and the biological relevance of the obtained results will be the use of infected cells as a starting material for purification of NendoU-containing complexes. The availability of highly specific antibodies that recognize NendoU-containing replicase subunits will also aid both approaches. This requirement, however, could be circumvented by the insertion of affinity tags in protein regions that are non-essential for RNA processing and virus viability. Such regions might be present in the family-specific N-terminal domains of coronavirus 15 and arterivirus nsp11. The available three-dimensional structures of two coronavirus nsp15 orthologs could guide the insertion of such tags<sup>6,29,60,105</sup>. Alternatively, random insertion mutagenesis could be employed to explore suitable insertion sites.

Isolation of NendoU-containing complexes may be facilitated by the association of the enzyme with membrane-bound viral replication complexes and the available protocols for their isolation from SARS-CoV- and EAV-infected cells<sup>77,98,99</sup>. This approach could not only aid the identification of interacting partners of NendoU-containing nsps, but will also help establish whether the *in vitro* substrate specificity of NendoU becomes narrower when in complex with other proteins. Alternatively, methods developed for the identification of *in vivo* RNA-protein interactions in cellular RNPs could be applied in the search for NendoU substrates. Briefly, these approaches rely on the irradiation of intact living cells or purified complexes with 254-nm UV light, which primarily cross-links proteins to RNA. The introduction of a covalent link between these two moieties permits the subsequent affinity purification of protein-RNA complexes under highly stringent conditions, ensuring that only specific interactions are detected<sup>21,87,88</sup>. After isolation of an RNP complex, cross-linked RNAs can be identified by RT-PCR and sequencing. Similar methods have been used successfully to map numerous RNA-protein interactions in small nucleolar RNPs<sup>21,88</sup>. Naturally, many potential pitfalls can be envisaged when trying to apply these methods to the identification of NendoU substrates, the most important of which would be the isolation of the enzyme in complex with its substrate or a cleavage product. The viability of NendoU active site mutants, together with their apparently



unimpaired ability to bind RNA molecules *in vitro* (Chapter 3 and <sup>6,30</sup>) may be exploited in pursuing this line of research.

In conclusion, our data on the arterivirus NendoU (Chapters 3 and 4) and that obtained by others on its coronavirus orthologs<sup>5-7,22,26,30,105</sup> provide a solid foundation for future research aimed at understanding the molecular details of NendoU functions during nidovirus replication. Apart from unique insights into the role of a replicative endoribonuclease in the life cycle of +RNA viruses, this knowledge will also be an important prerequisite for assessing the suitability of NendoU as a target for the development of antiviral compounds.

### **NSP1 IS A KEY FACTOR IN THE INTEGRAL COORDINATION OF THE ARTERIVIRUS REPLICATIVE CYCLE**

Nidovirus RNA synthesis in infected cells entails both amplification of the viral genome and the production of sg mRNAs, the latter serving as templates for the translation of viral structural and accessory protein genes. The process of sg mRNA synthesis is often referred to as “transcription”, to distinguish it from genome amplification (“replication”). The nidovirus genome and sg mRNAs are 3′-coterminal, while those of arteriviruses and coronaviruses also contain a common 5′“leader” sequence of 170-210 nucleotides (nt) or 55-92 nt, respectively<sup>34,76</sup>.

The majority of recent data on arteri- and coronavirus RNA synthesis supports a model proposed by Sawicki and Sawicki<sup>67</sup>, according to which the unique structure of arterivirus and coronavirus sg mRNAs derives from discontinuous extension during minus-strand RNA synthesis. This process resembles copy-choice RNA recombination and is guided by specific RNA signals termed transcription-regulatory sequences (TRS; core sequence 5′UCAACU 3′ in EAV). In the genomic RNA, these RNA motifs are found upstream of each structural protein gene in (body TRS) and in the 3′ end of the leader sequence (leader TRS) (recently reviewed in <sup>49,68</sup>). Minus-strand RNA synthesis is invariably initiated at the genomic 3′ end and can presumably be “attenuated” at a body TRS motif. The genomic leader TRS serves as a base-pairing target for the body TRS complement present at 3′ end of the nascent minus strand, which is translocated to the 5′-proximal region of the genomic template and extended with a copy of the leader sequence. A nested set of subgenome-length minus-strand templates are thus produced and serve as templates for the synthesis of the various sg mRNAs. Therefore, discontinuous minus-strand extension and the synthesis of a full-length complement of the genome likely “compete” for the same template – the genomic plus strand. This reasoning implies the existence of regulatory mechanisms that govern the temporal and quantitative control of the use of genomic RNA as a template for either full-length or subgenome-length minus-strand production.

## **EAV nsp1 links three key steps of the viral replicative cycle: replicase polyprotein processing, subgenomic mRNA production, and virion biogenesis**

Arterivirus nsp1 remains the sole nidovirus replicase subunit shown to be dispensable for genome amplification but essential for sg mRNA production. Based on the observation that accumulation of all sg mRNAs was blocked in the absence of nsp1, the protein was proposed to control a switch between replication and transcription via a predicted zinc finger (ZF) domain in its N-terminal region<sup>86</sup>. Apart from the ZF, nsp1 contains two additional conserved domains - papain-like cysteine protease alpha (PCP $\alpha$ ) and PCP $\beta$ , the former lacking proteolytic activity and the latter mediating nsp1 self-cleavage from pp1a/pp1ab<sup>11</sup>.

The data presented in Chapter 5 identified nsp1 as a multifunctional protein involved in three consecutive steps of the EAV replicative cycle: replicase protein proteolysis, sg mRNA production, and virion biogenesis. The autoproteolytic release of nsp1 from nascent replicase polyproteins was found to be required for EAV RNA synthesis (Chapter 5), even though replication can proceed in the absence of nsp1<sup>86</sup>. Impairing cleavage at the nsp1/2 site by the C-terminal PCP $\beta$  domain of nsp1 likely hinders the subsequent autoproteolytic cleavage of the nsp2/nsp3 site by a cysteine protease located in the N-terminal region of nsp2<sup>78</sup>. Accordingly, a mutation that blocked the self-release of nsp2 has been shown to dramatically affect downstream cleavage events in pp1a and result in a non-viable phenotype when analyzed in the EAV reverse genetics system<sup>55</sup>. Hence, the first proteolytic event in EAV replicase polyprotein processing – the autoproteolytic release of nsp1, may trigger a cascade of downstream proteolytic events that are essential for virus viability.

The postulated key role for the nsp1 ZF domain in transcription was supported by the complete and selective block of sg mRNA accumulation observed upon certain replacements of zinc-coordinating residues in the full-length EAV cDNA clone (Chapter 5). However, this domain is unlikely to be the sole determinant of the protein's function in transcription, because transcription-negative phenotypes were also obtained upon replacements of charged residues in the PCP $\alpha$  and PCP $\beta$  domains (Chapter 6). These results argue against the involvement of a distinct nsp1 subdomain in mediating sg mRNA production. The PCP $\beta$ -mediated autoproteolytic release of nsp1, on the other hand, does not seem to require an intact ZF domain (Chapters 5 and 7). Substitutions of zinc-coordinating residues, however, negatively impacted the solubility and presumably the folding of recombinant nsp1 expressed in *E. coli* (Chapter 7).

An intriguing phenotype was obtained upon replacements of residues in the nsp1 ZF and PCP $\alpha$  domains. These mutations had little effect on the accumulation of EAV mRNAs, but greatly decreased the production of infectious progeny particles, suggesting functional cooperation between the two domains in mediating a process required for efficient virion biogenesis (Chapters 5 and 6). It is still not known which part of the poorly characterized multistep process leading to the assembly and release of infectious EAV particles is perturbed by mutations in nsp1.



## **Nsp1 modulates minus-strand RNA accumulation to control the relative abundance of viral mRNAs**

The genomic RNA and sg mRNAs of arteri- and coronaviruses accumulate in specific relative molar ratios that are essentially constant until the peak of viral RNA synthesis has been reached, and it is largely unknown how these ratios are maintained<sup>2,13,64,66,69,71</sup>. Differential defects in the accumulation of EAV sg mRNAs have previously been observed upon mutagenesis of the core leader TRS and its flanking nucleotides. Altering the base-pairing potential between the leader TRS and the body TRS complements in nascent minus strands through mutation of either element showed that the size and stability of this duplex region is an important determinant of the levels to which the various sg mRNAs accumulate<sup>51,52</sup>. The data presented in Chapter 6 revealed for the first time that mRNA-specific modulation of viral RNA levels in EAV-infected cells can be brought about by mutations in a protein factor: nsp1. These results, combined with the data presented in Chapter 5, provide genetic evidence that balanced accumulation of EAV mRNA species is also regulated by nsp1, implying the existence of a functional link between this replicase subunit and the TRS network of RNA signals. Unlike the leader TRS, however, nsp1 is also involved in a process controlling the amount of viral genome that accumulates in EAV-infected cells.

The relative molar ratios of coronavirus mRNAs and their corresponding minus-strand templates are similar, implying that coronavirus mRNA abundance is primarily determined at the level of minus-strand synthesis<sup>2,64,71</sup>. The quantitative analysis of individual minus-strand RNAs produced in nidovirus-infected cells, however, is inherently problematic due to their colinearity and low abundance, and is further complicated by the presence of a large excess (> 100-fold) of plus strands<sup>65,101</sup>. To meet these challenges, we developed a sensitive ribonuclease (RNase) protection assay for the detection and accurate quantitation of genome- and subgenome-length EAV minus-strand RNAs (Chapter 6). This assay permits the first-cycle analysis of minus-strand accumulation, because it requires a relatively small number of EAV-transfected cells, and is therefore well-suited for the characterization of minus-strand production by EAV mutants defective in infectious progeny production. This technical advance allowed us to establish that nsp1 maintains the balance among the seven EAV mRNAs, likely by controlling the levels of their corresponding minus-strand templates (Chapter 6).

Expanding the current set of methods available for the analysis of EAV mutant phenotypes with a protocol for the quantitation of viral minus-strand RNAs will undoubtedly assist in pinpointing the mechanisms underlying defects in viral mRNA production. For example, the protocol could be used to determine whether substitutions in EAV NendoU that negatively impact viral mRNA abundance (Chapter 3) have comparable effects on minus-strand accumulation and thus help establish whether NendoU functions during plus-strand or minus-strand synthesis. Analysis of the effects of well-characterized leader TRS and body TRS mutations<sup>51,52,92</sup> on subgenome-length minus strand accumulation would also facilitate the characterization of the roles these RNA motifs play in the

discontinuous extension of minus strands. A better understanding of this latter process could be facilitated by the characterization of postulated transitory intermediates, such as “antileaderless” minus strands that may accumulate when the transfer of the nascent minus strand to the leader TRS region is impaired<sup>51,52,91,92</sup>.

Using the RNase protection assay described in Chapter 6, we attempted to detect such intermediates by hybridization of total RNA from EAV-infected cells with a probe complementary to the region of the EAV genomic minus strand containing the body TRS complement of RNA2. Theoretically, this probe should protect two fragments: a larger fragment derived from the genomic minus strand, and a smaller fragment derived from the body sequence complement present uniquely in the subgenome-length minus-strand template of RNA2 [(-)RNA2]. This approach, however, consistently yielded only a single protected fragment originating from the genomic minus strand when samples from EAV-infected cells were analyzed (data not shown). Our inability to detect a (-) RNA2-derived fragment in this experimental setup may be related to the lower relative abundance of RNA2 (and its minus-strand template) when compared to the genomic RNA (see Chapter 6). An entirely different experimental approach, such as e. g. RNA ligase-mediated amplification of cDNA ends<sup>38</sup> may be better suited to establish whether antileaderless subgenomic minus strands accumulate as a result of mutations in regulatory proteins or RNA sequences.

### **What are the molecular mechanisms of nsp1 function in EAV RNA synthesis?**

There is ample genetic evidence for the critical role of EAV nsp1 in regulating the balance between replication and transcription, as well as in controlling the relative abundance of viral mRNAs (<sup>86</sup> and Chapters 5 and 6). Examining the interactions of nsp1 with other viral proteins and RNA motifs critically involved in viral RNA synthesis, and discontinuous minus-strand extension in particular, is an obvious starting point in dissecting the molecular bases of the various regulatory activities of nsp1 in the EAV replicative cycle.

It will be important to establish whether nsp1 can associate with nsp9 and nsp10, the EAV RdRp and helicase, respectively. A single amino acid substitution in nsp10 can selectively reduce the accumulation levels of subgenome-length plus and minus strands by ~500-fold without affecting the *in vitro* duplex-unwinding activity of recombinant nsp10<sup>72,95,101</sup>. Perturbation of an nsp1-nsp10 interaction required for the discontinuous extension of minus strands would certainly be an appealing explanation for the transcription defect of the nsp10 mutant. Our attempts to isolate an nsp1-nsp10 or an nsp1-nsp9 protein complex in EAV-infected cells by co-immunoprecipitation under various experimental conditions have not been successful so far (data not shown), a result that may reflect the dynamic nature of these putative complexes. Experimental verification of their formation could be pursued by examining the ability of these proteins to associate *in vitro*, which will be facilitated by the available protocols for the purification of recombinant nsp1 (Chapter 7), nsp9<sup>3</sup>, and nsp10<sup>73</sup> from *E. coli*. Ultimately, an *in vitro* assay reconstituting RdRp attenuation during minus-strand RNA synthesis should be devel-

oped to assess the effects of nsp1 and nsp10 on this process. To this end, the currently available EAV RdRp assay<sup>3</sup> must be advanced to support robust *in vitro* RdRp activity of recombinant EAV nsp9 on viral templates.

In the discontinuous minus-strand synthesis model, the relative abundance of sg mRNAs is presumably determined by the “attenuation rate” at each of the successive body TRS motifs encountered during minus-strand synthesis, and the data we obtained in Chapter 6 can be reconciled with this model. Interactions of nsp1 with body TRS motifs in the plus strand or their complements in nascent minus strands are therefore obvious candidates for promoting viral RdRp pausing and/or nascent strand transfer to the genomic leader region. Our preliminary data suggest that recombinant nsp1 can interact with RNA both as a monomer and a dimer, and exhibits low binding specificity *in vitro* (A.J.W te Velthuis and D.D. Nedialkova, unpublished observations). Ongoing experiments are aimed at establishing whether nsp1 can operate in a sequence-specific manner, by binding to RNA sequences in EAV body TRS regions or their minus-strand complements, and discriminate between the various TRS motifs. Considering the apparent ability of nsp1 to bind RNA in a dimeric form, it will also be interesting to determine if the protein can aid the juxtaposition of sequence elements in the leader TRS hairpin and the nascent minus strand, and thereby facilitate base-pairing between the leader TRS and the body TRS complement. This base-pairing step presumably ensures the fidelity of strand transfer during discontinuous minus-strand synthesis, and the core TRS sequence is thus unlikely a (major) determinant of sequence-specific body TRS recognition by protein factors<sup>51,92,100</sup>. In line with this notion, a construct carrying five mutations (5'-UCAACU-3' - 5'-AGUUGU-3') in the EAV leader TRS and RNA7 body TRS still permitted a certain level of RNA7 accumulation<sup>100</sup>. Also, a recombinant SARS-CoV with a “rewired transcription circuit”, carrying a 3-nt substitution in the core sequences of the leader and all body TRSs, replicates efficiently in cell culture<sup>106</sup>. Hence, body TRS motifs are likely recognized based on higher-order RNA structure present in the genomic RNA template or the nascent minus strand, rather than in a (core) sequence-dependent manner. This observation should be taken into account when examining the ability of recombinant nsp1 to bind these motifs *in vitro*.

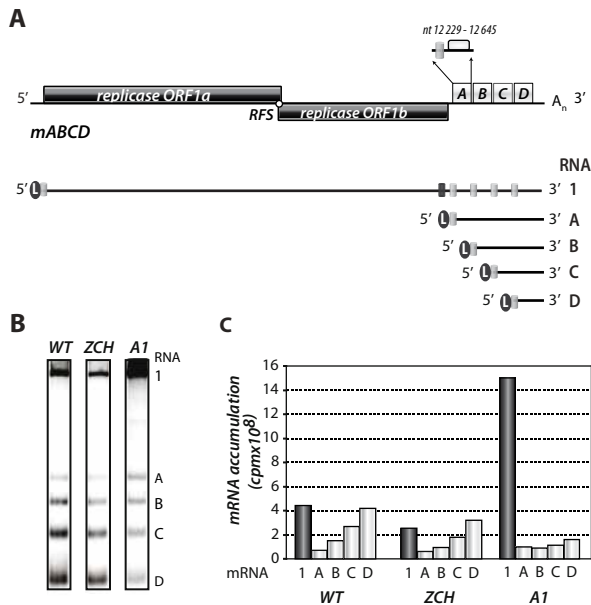
The unique sequence context of each body TRS motif and its proximity to the genomic 3' end can also influence the accumulation of cognate sg mRNA species in EAV-infected cells. This was established in a prior study with the use of an EAV replicon derived from the full-length clone. In this construct, the sequence context of a body TRS motif was standardized by inserting several copies of an RNA7 body TRS cassette directly downstream of the EAV replicase gene in a head-to-tail fashion<sup>50</sup>. This cassette contained the core body TRS sequence of RNA7, flanked by 21 nt of natural upstream and 388 nt natural downstream sequences, including the complete EAV N protein gene. In this setting, all but the 5'-proximal TRS motifs have an identical sequence context over more than 400 nt, but are found at different distances from the genomic 3' end (Fig. 2A). The number of sg mRNAs produced in cells transfected with this construct corresponded to the number of (functional) body TRS motifs. Smaller sg mRNAs accumulated to higher levels in this





setting, resulting in a gradient of sg mRNA abundance<sup>50</sup>. Such “polar attenuation” may reflect the shorter time period needed for the synthesis of smaller RNAs. Alternatively, it can result from the association of a protein factor with the RNA-synthesizing complex necessary for commitment of the latter to discontinuous minus-strand extension. If this factor is present in a limiting amount or mediates a rate-limiting step, attenuation of minus-strand synthesis would occur more frequently at 3′-proximal body TRS regions<sup>41,50</sup>.

In an attempt to assess whether mRNA-specific defects associated with nsp1 mutations were related to the proximity of body TRS motifs to the genomic 3′ end, we engineered the ZCH and A1 mutations (Chapter 6) in the construct containing four repeats



**Figure 2. Effects of nsp1 mutations on the relative abundance of sg mRNAs produced from a construct with repeated RNA7 body TRS cassettes.** (A) Schematic representation of the mABCD construct described by Pasternak *et al.*<sup>50</sup>. Four repeats of a sequence cassette containing 417nt from the EAV genome (nt 12,229 to 12,645) were inserted in a head-to-tail fashion immediately downstream of the EAV replicase gene, replacing the structural protein ORFs. The cassette includes the body TRS of RNA7 (indicated with a grey rectangle), flanked by 21 nt and 388 nt of natural upstream and downstream sequences, respectively. Consequently, this cassette contains the entire ORF coding for N protein. Transfection of BHK-21 cells with RNA transcribed *in vitro* from the mABCD construct results in the amplification of the genomic RNA (RNA1) and the synthesis of four sg mRNAs (RNAs A – D), schematically drawn in the lower panel. The body TRS motif of RNA2, located in replicase ORF1b (shown with a black rectangle), was inactivated by two nucleotide substitutions. (B) The ZCH and A1 mutations in nsp1 (see chapter 6) were introduced into the mABCD background, and BHK-21 cells were transfected with RNA transcribed from wild-type or mutant mABCD derivatives. Total intracellular RNA was isolated at 11 h post-transfection and resolved by denaturing formaldehyde electrophoresis. EAV-specific mRNAs were detected by hybridization of the gel with a <sup>32</sup>P-labelled probe complementary to the 3′-end of the viral genome. The positions of the genomic RNA and the four sg mRNAs are indicated. (C) Accumulation levels of each viral mRNA in panel (B) were quantified by phosphorimaging. Dark grey bars denote genomic RNA levels.

of the RNA7 body TRS cassette (mABCD) described by Pasternak *et al.* (Fig. 2A). In the background of the full-length EAV cDNA clone, the ZCH mutation resulted in ~2-fold upregulation of RNA7 accumulation (Chapter 6), but had little effect on the abundance of sg mRNAs when introduced in the mABCD construct (Fig. 2, B and C). The observed background-specific difference must be linked to the sequence context of TRS elements. For instance, an increase in sg mRNA accumulation may require a larger portion of the upstream sequence flanking the body TRS of RNA7 than the 21 nt present in each cassette of the mABCD construct. In addition, the size gradient of sg mRNA abundance was not affected by the ZCH mutation. By contrast, when the A1 mutation was introduced in the mABCD background, all sg mRNAs accumulated to similar levels, irrespective of their sizes (Fig; 2, B and C). Also, a pronounced enhancement of genomic RNA accumulation was seen in the A1 mutant, similar to what we had previously observed in the background of the full-length EAV cDNA clone (Chapter 6). The basis for the virtual absence of “polar attenuation” in the mABCD construct that contains the A1 mutation in nsp1 is not immediately evident. It may be associated with the increased availability of viral nonstructural proteins and/or template, both stemming from the higher levels of genomic RNA accumulation. A role of viral replicase subunits (and nsp1 in particular) as limiting factors in EAV sg mRNA synthesis is thus a tempting hypothesis. In order to assess whether nsp1 availability is important for sg mRNA accumulation, however, it will be necessary to uncouple the effects of nsp1 mutations on genomic RNA accumulation from those on sg mRNA abundance. This may be addressed with the use of an inducible cell line, in which nsp1 is supplied *in trans* and its expression levels can be quantitatively modulated. In this setting, the effects of nsp1 abundance on replication and transcription of an EAV mutant lacking nsp1 can be examined. Our efforts to generate nsp1-expressing stable cell lines have unfortunately been unsuccessful so far, probably due to the high toxicity of the heterologously expressed protein (data not shown).

Finally, it should be stressed that evidence for the involvement of a replicase protein in the mRNA-specific control of nidovirus RNA synthesis has only been obtained in EAV so far. Uncoupling of genome replication from sg mRNA production through mutations in nonstructural proteins has also only been observed upon substitutions in nsp1 of EAV and PRRSV (Chapters 5 and 6, <sup>33,86</sup>), as well as EAV nsp10<sup>95,101</sup>. It is therefore unclear whether the synthesis of full-length and subgenome-length RNAs by members of other nidovirus families also require protein complexes with a different composition. Alternatively, the viral RdRp-containing replicase subunit of *Coronaviridae* and *Roniviridae* representatives can contain domains responsible for the recognition of TRS motifs, especially in view of the unprecedented size of this protein among viral RdRps<sup>20,85</sup>. Mechanistic differences in replication and transcription among the different nidovirus groups certainly exist, as illustrated by the lack of a common leader sequence in sg mRNAs of some nidovirus representatives<sup>10,75,102</sup>. The larger genome size and larger set of replicative proteins encoded by viruses belonging to the *Coronaviridae* and *Roniviridae* families<sup>18,26,42,74</sup> are also likely to be manifested in differences among the molecular mechanisms of viral RNA synthesis.

Nevertheless, the fundamental aspects of this process are likely shared among nidoviruses, considering the evolutionary relationship among their key replicative enzymes.

### **The relative abundance of EAV mRNAs is fine-tuned to allow efficient virus production**

The data presented in Chapter 6 allowed us to postulate the existence of a previously unknown link between the regulation of individual nidovirus mRNA levels and the efficiency of infectious progeny production. Mutations in nsp1 that exerted differential effects on EAV mRNA and corresponding protein abundance also adversely affected the assembly of infectious virions. We isolated numerous nsp1 pseudorevertants with compensatory mutations which invariably rescued both efficient virus production and balanced EAV mRNA accumulation. A closer inspection of the relationship between infectious virus yield and viral mRNA accumulation revealed that the magnitude of imbalance between different mRNA species is a principal factor affecting progeny yield in EAV; in fact, there seems to be a remarkably linear relationship between virus titers and mRNA species imbalance (Chapter 6). Collectively, these results reveal that mRNA-specific modulation of viral RNA levels is a fundamental mechanism for the quantitative control of viral protein expression, which has important implications for the late stages of the EAV replicative cycle. Our results also suggest that nsp1 promotes virion biogenesis by acting at different stages of the EAV replicative cycle – one connected to the modulation of viral mRNA accumulation, as well as an additional, currently unknown step downstream of viral RNA synthesis (Chapters 5 and 6).

Altering relative abundances of viral structural proteins by changing the levels of their mRNA templates could negatively impact virion assembly by several mechanisms. For example, the stoichiometry of protein complexes which drive this process and/or confer infectivity to progeny particles can be altered. The decreased ratio between structural proteins and genomic RNA we observed in some nsp1 mutants may also adversely affect infectious progeny production. Accordingly, virion preparations from nsp1 mutants with imbalanced mRNA accumulation profiles contained not only fewer virus particles, but also had lower specific infectivity (Chapter 6). It will be interesting to determine whether virions produced by nsp1 mutants with defects in progeny production but not in viral RNA accumulation are also characterized by low specific infectivity. Elucidating the molecular basis of nsp1 function in virion biogenesis, however, requires more comprehensive knowledge of the molecular processes that govern genome encapsidation, as well as the assembly and secretion of infectious progeny particles from EAV-infected cells.



## UNDERSTANDING NIDOVIRUS REPLICATION: A CASE FOR REDUCTIONISM

A widely used approach in research on complex biological systems is reducing them to a set of fundamental parts, the functioning and interactions of which are subsequently characterized separately. This method has been successfully applied in studies on the replicative cycle of numerous +RNA viruses. In particular, various strategies have been used to uncouple genome replication from its translation and encapsidation in order to examine the critical determinants for each process (see Chapter 2). Similar approaches could be applied to address some of the fundamental questions about the regulation of key steps in the nidovirus replicative cycle, although experimental setups may have to be tailored to accommodate the complexity of nidovirus replication.

In analogy to research conducted with DENV and YFV<sup>1,28</sup>, nidovirus replicons encoding reporter proteins with enzymatic activities which can be monitored with high sensitivity (such as firefly luciferase), could be designed to discriminate between mutations affecting viral genome translation or RNA synthesis. Expression of green fluorescent protein from the replicase gene of EAV without major effects on viral replication has already been reported, demonstrating that insertion of foreign (reporter) genes in certain regions of the EAV replicase is feasible<sup>93</sup>. Also, generation of nidovirus replicon RNAs lacking structural protein genes from DNA-based vectors, combined with *trans*-complementation approaches, should help establish whether replication of nidovirus genomes is a prerequisite for their incorporation into viral particles, as shown for KUN virus and PV<sup>31,46</sup>. Constructs containing the EAV replicase gene-coding sequence downstream of the T7 RNA polymerase promoter are already available and have been instrumental in studies on EAV polyprotein processing in the recombinant vaccinia virus/T7 RNA polymerase system<sup>79,90,96,104</sup>. This experimental approach has permitted the analysis of proteolysis defects associated with mutations in the main EAV protease that are lethal in the context of the complete EAV replicative cycle. EAV genome translation was thus effectively uncoupled from its amplification in this system, which may serve as a basis for future experiments on the determinants of these two key viral processes.

Apart from genetic approaches, development of a cell-free assay for nidovirus RNA synthesis, analogous to the one available for PV<sup>43</sup>, would provide an invaluable tool for characterization of mutations that block nidovirus replication early in infection. Ultimately, (a combination of) methods should permit the identification of specific steps in genome translation and replication affected by mutations that are now classified simply as “nonviable” with conventional reverse genetics. The recently developed assays for the isolation of active viral replication/transcription complexes (RTCs) from SARS-CoV- and EAV-infected cells<sup>98,99</sup>, however, can currently only be used to analyze wild-type RTCs, or mutant RTCs that support robust RNA synthesis and virus production. Nevertheless, the availability of purified active nidovirus RTCs will greatly benefit analysis of the composition of the nidovirus RNA-synthesizing machinery, as well as identification of the host factors shown to be required for nidovirus RC activity *in vitro*. Ideally, the question of whether genome-length and subgenome-length are produced by RdRp complexes of

different composition, i.e. the existence of a separate “replicase” and a “transcriptase”, could also be addressed by attempting to purify the two activities, as has been achieved for vesicular stomatitis virus, a minus-strand RNA virus<sup>57</sup>. This, however, will likely require the isolation of complexes that are actively engaged in minus-strand synthesis, and it is unclear whether nidovirus RTCs isolated according to the currently available protocols can mediate production of minus strands<sup>98,99</sup>.

Alternative “bottom-up” approaches relying on recently developed *in vitro* RdRp activity assays<sup>3,85</sup> can also be envisaged, where known components (e.g. membranes, recombinant nonstructural proteins, viral RNA templates) can be combined to reconstruct active RNA-synthesizing complexes exhibiting at least some of the properties of nidovirus RTCs. The large number of nidovirus replicase subunits and the potential critical roles of polyprotein processing intermediates in viral RNA synthesis, however, may complicate such approaches. Taking into account also the large size of nidovirus genomic RNAs (13–32 kb) and the mounting evidence that long-range RNA-RNA interactions are key players in regulation of the +RNA virus replicative cycle, reconstruction of nidovirus RNA-synthesizing complexes from their individual components will certainly be a formidable task.



## REFERENCES

1. **Alvarez, D. E., A. L. De Lella Ezcurra, S. Fucito, and A. V. Gamarnik.** 2005. Role of RNA structures present at the 3'UTR of dengue virus on translation, RNA synthesis, and viral replication. *Virology* **339**:200-212
2. **Baric, R. S. and B. Yount.** 2000. Subgenomic negative-strand RNA function during mouse hepatitis virus infection. *J.Virol.* **74**:4039-4046
3. **Beerens, N., B. Selisko, S. Ricagno, I. Imbert, L. van der Zanden, E. J. Snijder, and B. Canard.** 2007. De novo initiation of RNA synthesis by the arterivirus RNA-dependent RNA polymerase. *J.Virol.* **81**:8384-8395
4. **Behrens, S. E., L. Tomei, and R. De Francesco.** 1996. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO J.* **15**:12-22
5. **Bhardwaj, K., L. Guarino, and C. C. Kao.** 2004. The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. *J.Virol.* **78**:12218-12224
6. **Bhardwaj, K., S. Palaninathan, J. M. Alcantara, L. L. Yi, L. Guarino, J. C. Sacchettini, and C. C. Kao.** 2008. Structural and functional analyses of the severe acute respiratory syndrome coronavirus endoribonuclease Nsp15. *J.Biol.Chem.* **283**:3655-3664
7. **Bhardwaj, K., J. Sun, A. Holzenburg, L. A. Guarino, and C. C. Kao.** 2006. RNA recognition and cleavage by the SARS coronavirus endoribonuclease. *J.Mol.Biol.* **361**:243-256
8. **Carpousis, A. J.** 2007. The RNA degradosome of *Escherichia coli*: an mRNA-degrading machine assembled on RNase E. *Annu.Rev.Microbiol.* **61**:71-87
9. **Chen, Y., H. Cai, J. Pan, N. Xiang, P. Tien, T. Ahola, and D. Guo.** 2009. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc.Natl. Acad.Sci.U.S.A* **106**:3484-3489
10. **Cowley, J. A., C. M. Dimmock, and P. J. Walker.** 2002. Gill-associated nidovirus of *Penaeus monodon* prawns transcribes 3'-coterminal subgenomic mRNAs that do not possess 5'-leader sequences. *J.Gen.Virol.* **83**:927-935
11. **den Boon, J. A., K. S. Faaberg, J. J. Meulenberg, A. L. Wassenaar, P. G. Plagemann, A. E. Gorbalenya, and E. J. Snijder.** 1995. Processing and evolution of the N-terminal region of the arterivirus replicase ORF1a protein: identification of two papainlike cysteine proteases. *J.Virol.* **69**:4500-4505
12. **den Boon, J. A., E. J. Snijder, E. D. Chirnside, A. A. de Vries, M. C. Horzinek, and W. J. Spaan.** 1991. Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. *J.Virol.* **65**:2910-2920
13. **den Boon, J. A., W. J. M. Spaan, and E. J. Snijder.** 1995. Equine arteritis virus subgenomic RNA transcription - UV inactivation and translation inhibition studies. *Virology* **213**:364-372
14. **Durand, S., G. Richard, M. Bisaglia, S. Laalami, F. Bontems, and M. Uzan.** 2006. Activation of RegB endoribonuclease by S1 ribosomal protein requires an 11 nt conserved sequence. *Nucleic Acids Res.* **34**:6549-6560
15. **Eigen, M.** 1993. The origin of genetic information: viruses as models. *Gene* **135**:37-47
16. **Eriksson, K. K., L. Cervantes-Barragan, B. Ludewig, and V. Thiel.** 2008. Mouse hepatitis virus liver pathology is dependent on ADP-ribose-1"-phosphatase, a viral function conserved in the alpha-like supergroup. *J.Virol.* **82**:12325-12334

17. **Ferrari, E., J. Wright-Minogue, J. W. Fang, B. M. Baroudy, J. Y. Lau, and Z. Hong.** 1999. Characterization of soluble hepatitis C virus RNA-dependent RNA polymerase expressed in *Escherichia coli*. *J.Virol.* **73**:1649-1654
18. **Gorbalenya, A. E.** 2001. Big nidovirus genome. When count and order of domains matter. *Adv. Exp.Med.Biol.* **494**:1-17
19. **Gorbalenya, A. E., L. Enjuanes, J. Ziebuhr, and E. J. Snijder.** 2006. Nidovirales: evolving the largest RNA virus genome. *Virus Res.* **117**:17-37
20. **Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov.** 1989. Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Res.* **17**:4847-4861
21. **Granneman, S., G. Kudla, E. Petfalski, and D. Tollervey.** 2009. Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc.Natl.Acad.Sci.U.S.A* **106**:9613-9618
22. **Guarino, L. A., K. Bhardwaj, W. Dong, J. Sun, A. Holzenburg, and C. Kao.** 2005. Mutational analysis of the SARS virus Nsp15 endoribonuclease: identification of residues affecting hexamer formation. *J.Mol.Biol.* **353**:1106-1117
23. **Hausmann, Y., G. Roman-Sosa, H. J. Thiel, and T. Rumenapf.** 2004. Classical swine fever virus glycoprotein E rns is an endoribonuclease with an unusual base specificity. *J.Virol.* **78**:5507-5512
24. **Imbert, I., J. C. Guillemot, J. M. Bourhis, C. Bussetta, B. Coutard, M. P. Egloff, F. Ferron, A. E. Gorbalenya, and B. Canard.** 2006. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J.* **25**:4933-4942
25. **Imbert, I., E. J. Snijder, M. Dimitrova, J. C. Guillemot, P. Lecine, and B. Canard.** 2008. The SARS-Coronavirus PLnc domain of nsp3 as a replication/transcription scaffolding protein. *Virus Res.* **133**:136-148
26. **Ivanov, K. A., T. Hertzog, M. Rozanov, S. Bayer, V. Thiel, A. E. Gorbalenya, and J. Ziebuhr.** 2004. Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc.Natl.Acad. Sci.U.S.A* **101**:12694-12699
27. **Jeffares, D. C., A. M. Poole, and D. Penny.** 1998. Relics from the RNA world. *J.Mol.Evol.* **46**:18-36
28. **Jones, C. T., C. G. Patkar, and R. J. Kuhn.** 2005. Construction and applications of yellow fever virus replicons. *Virology* **331**:247-259
29. **Joseph, J. S., K. S. Saikatendu, V. Subramanian, B. W. Neuman, M. J. Buchmeier, R. C. Stevens, and P. Kuhn.** 2007. Crystal structure of a monomeric form of severe acute respiratory syndrome coronavirus endonuclease nsp15 suggests a role for hexamerization as an allosteric switch. *J.Virol.* **81**:6700-6708
30. **Kang, H., K. Bhardwaj, Y. Li, S. Palaninathan, J. Sacchettini, L. Guarino, J. L. Leibowitz, and C. C. Kao.** 2007. Biochemical and genetic analyses of murine hepatitis virus Nsp15 endoribonuclease. *J.Virol.* **81**:13587-13597
31. **Khromykh, A. A., A. N. Varnavski, P. L. Sedlak, and E. G. Westaway.** 2001. Coupling between replication and packaging of flavivirus RNA: evidence derived from the use of DNA-based full-length cDNA clones of Kunjin virus. *J.Virol.* **75**:4633-4640
32. **Knoops, K., M. Kikkert, S. H. Worm, J. C. Zevenhoven-Dobbe, Y. van der Meer, A. J. Koster, A. M. Mommaas, and E. J. Snijder.** 2008. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS.Biol.* **6**:e226
33. **Kroese, M. V., J. C. Zevenhoven-Dobbe, J. N. Bos-de Ruijter, B. P. Peeters, J. J. Meulenber, L. A. Cornelissen, and E. J. Snijder.** 2008. The nsp1alpha and nsp1 papain-like autoproteases are essential for porcine reproductive and respiratory syndrome virus RNA synthesis. *J.Gen.Virol.* **89**:494-499



34. **Lai, M. M. and D. Cavanagh.** 1997. The molecular biology of coronaviruses. *Adv.Virus Res.* **48**:1-100
35. **Lebars, I., R. M. Hu, J. Y. Lallemand, M. Uzan, and F. Bontems.** 2001. Role of the substrate conformation and of the S1 protein in the cleavage efficiency of the T4 endoribonuclease RegB. *J.Biol. Chem.* **276**:13264-13272
36. **Lee, K., X. Zhan, J. Gao, J. Qiu, Y. Feng, R. Meganathan, S. N. Cohen, and G. Georgiou.** 2003. RraA, a protein inhibitor of RNase E activity that globally modulates RNA abundance in *E. coli*. *Cell* **114**:623-634
37. **Li, W. M., T. Barnes, and C. H. Lee.** 2009. Endoribonucleases - enzymes gaining spotlight in mRNA metabolism. *FEBS Journal* **277**:627-641
38. **Liu, X. and M. A. Gorovsky.** 1993. Mapping the 5' and 3' ends of *Tetrahymena thermophila* mRNAs using RNA ligase mediated amplification of cDNA ends (RLM-RACE). *Nucleic Acids Res.* **21**:4954-4960
39. **Magkouras, I., P. Matzener, T. Rumenapf, E. Peterhans, and M. Schweizer.** 2008. RNase-dependent inhibition of extracellular, but not intracellular, dsRNA-induced interferon synthesis by Erns of pestiviruses. *J.Gen.Virol.* **89**:2501-2506
40. **Meyers, G., A. Saalmuller, and M. Buttner.** 1999. Mutations abrogating the RNase activity in glycoprotein E(rns) of the pestivirus classical swine fever virus lead to virus attenuation. *J.Virol.* **73**:10224-10235
41. **Miller, W. A. and G. Koev.** 2000. Synthesis of subgenomic RNAs by positive-strand RNA viruses. *Virology* **273**:1-8
42. **Minskaia, E., T. Hertzog, A. E. Gorbalenya, V. Campanacci, C. Cambillau, B. Canard, and J. Ziebuhr.** 2006. Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc.Natl.Acad.Sci.U.S.A* **103**:5108-5113
43. **Molla, A., A. V. Paul, and E. Wimmer.** 1991. Cell-free, de novo synthesis of poliovirus. *Science* **254**:1647-1651
44. **Morl, M. and A. Marchfelder.** 2001. The final cut. The importance of tRNA 3'-processing. *EMBO Rep.* **2**:17-20
45. **Nogues, M. V., M. Vilanova, and C. M. Cuchillo.** 1995. Bovine pancreatic ribonuclease A as a model of an enzyme with multiple substrate binding sites. *Biochim.Biophys.Acta* **1253**:16-24
46. **Nugent, C. I., K. L. Johnson, P. Sarnow, and K. Kirkegaard.** 1999. Functional coupling between replication and packaging of poliovirus replicon RNA. *J.Virol.* **73**:427-435
47. **Pan, J., X. Peng, Y. Gao, Z. Li, X. Lu, Y. Chen, M. Ishaq, D. Liu, M. L. Dediego, L. Enjuanes, and D. Guo.** 2008. Genome-wide analysis of protein-protein interactions and involvement of viral proteins in SARS-CoV replication. *PLoS.One.* **3**:e3299
48. **Pasternak, A. O., A. P. Gulyaev, W. J. M. Spaan, and E. J. Snijder.** 2000. Genetic manipulation of arterivirus alternative mRNA leader-body junction sites reveals tight regulation of structural protein expression. *J. Virol.* **74**:11642-11653
49. **Pasternak, A. O., W. J. Spaan, and E. J. Snijder.** 2006. Nidovirus transcription: how to make sense...? *J.Gen.Virol.* **87**:1403-1421
50. **Pasternak, A. O., W. J. M. Spaan, and E. J. Snijder.** 2004. Regulation of relative abundance of arterivirus subgenomic mRNAs. *J. Virol.* **78**:8102-8113
51. **Pasternak, A. O., E. van den Born, W. J. M. Spaan, and E. J. Snijder.** 2001. Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. *EMBO J.* **20**:7220-7228



52. **Pasternak, A. O., E. van den Born, W. J. M. Spaan, and E. J. Snijder.** 2003. The stability of the duplex between sense and antisense transcription-regulating sequences is a crucial factor in arterivirus subgenomic mRNA synthesis. *J.Virol.* **77**:1175-1183
53. **Pedersen, K. W., Y. van der Meer, N. Roos, and E. J. Snijder.** 1999. 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.* **73**:2016-2026
54. **Plotch, S. J., O. Palant, and Y. Gluzman.** 1989. Purification and properties of poliovirus RNA polymerase expressed in *Escherichia coli*. *J.Virol.* **63**:216-225
55. **Posthuma, C. C., K. W. Pedersen, Z. Lu, R. G. Joosten, N. Roos, J. C. Zevenhoven-Dobbe, and E. J. Snijder.** 2008. Formation of the arterivirus replication/transcription complex: a key role for nonstructural protein 3 in the remodeling of intracellular membranes. *J.Virol.* **82**:4480-4491
56. **Prentice, E., J. McAuliffe, X. Lu, K. Subbarao, and M. R. Denison.** 2004. Identification and characterization of severe acute respiratory syndrome coronavirus replicase proteins. *J.Virol.* **78**:9977-9986
57. **Qanungo, K. R., D. Shaji, M. Mathur, and A. K. Banerjee.** 2004. Two RNA polymerase complexes from vesicular stomatitis virus-infected cells that carry out transcription and replication of genome RNA. *Proc.Natl.Acad.Sci.U.S.A* **101**:5952-5957
58. **Raines, R. T.** 1998. Ribonuclease A. *Chem.Rev.* **98**:1045-1066
59. **Renzi, F., E. Caffarelli, P. Laneve, I. Bozzoni, M. Brunori, and B. Vallone.** 2006. The structure of the endoribonuclease XendoU: From small nucleolar RNA processing to severe acute respiratory syndrome coronavirus replication. *Proc.Natl.Acad.Sci.U.S.A* **103**:12365-12370
60. **Ricagno, S., M. P. Egloff, R. Ulferts, B. Coutard, D. Nurizzo, V. Campanacci, C. Cambillau, J. Ziebuhr, and B. Canard.** 2006. Crystal structure and mechanistic determinants of SARS coronavirus nonstructural protein 15 define an endoribonuclease family. *Proc.Natl.Acad.Sci.U.S.A* **103**:11892-11897
61. **Rohayem, J., K. Jager, I. Robel, U. Scheffler, A. Temme, and W. Rudolph.** 2006. Characterization of norovirus 3Dpol RNA-dependent RNA polymerase activity and initiation of RNA synthesis. *J.Gen.Virol.* **87**:2621-2630
62. **Roth-Cross, J. K., H. Stokes, G. Chang, M. M. Chua, V. Thiel, S. R. Weiss, A. E. Gorbalenya, and S. G. Siddell.** 2009. Organ-specific attenuation of murine hepatitis virus strain A59 by replacement of catalytic residues in the putative viral cyclic phosphodiesterase ns2. *J.Virol.* **83**:3743-3753
63. **Ruckman, J., S. Ringquist, E. Brody, and L. Gold.** 1994. The bacteriophage T4 regB ribonuclease. Stimulation of the purified enzyme by ribosomal protein S1. *J.Biol.Chem.* **269**:26655-26662
64. **Sawicki, D. L., T. Wang, and S. G. Sawicki.** 2001. The RNA structures engaged in replication and transcription of the A59 strain of mouse hepatitis virus. *J.Gen.Virol.* **82**:385-396
65. **Sawicki, S. G. and D. L. Sawicki.** 1986. Coronavirus minus-strand RNA synthesis and effect of cycloheximide on coronavirus RNA synthesis. *J.Virol.* **57**:328-334
66. **Sawicki, S. G. and D. L. Sawicki.** 1990. Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. *J.Virol.* **64**:1050-1056
67. **Sawicki, S. G. and D. L. Sawicki.** 1995. Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. *Adv.Exp.Med.Biol.* **380**:499-506
68. **Sawicki, S. G., D. L. Sawicki, and S. G. Siddell.** 2007. A contemporary view of coronavirus transcription. *J.Virol.* **81**:20-29
69. **Schaad, M. C. and R. S. Baric.** 1994. Genetics of mouse hepatitis virus transcription: evidence that subgenomic negative strands are functional templates. *J.Virol.* **68**:8169-8179
70. **Schneider, R., G. Unger, R. Stark, E. Schneider-Scherzer, and H. J. Thiel.** 1993. Identification of a structural glycoprotein of an RNA virus as a ribonuclease. *Science* **261**:1169-1171





71. **Sethna, P. B., S. L. Hung, and D. A. Brian.** 1989. Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. *Proc.Natl.Acad.Sci.U.S.A* **86**:5626-5630
72. **Seybert, A., C. C. Posthuma, L. C. van Dinten, E. J. Snijder, A. E. Gorbalenya, and J. Ziebuhr.** 2005. A complex zinc finger controls the enzymatic activities of nidovirus helicases. *J.Virol.* **79**:696-704
73. **Seybert, A., L. C. van Dinten, E. J. Snijder, and J. Ziebuhr.** 2000. Biochemical characterization of the equine arteritis virus helicase suggests a close functional relationship between arterivirus and coronavirus helicases. *J. Virol.* **74**:9586-9593
74. **Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. Poon, Y. Guan, M. Roza-nov, W. J. Spaan, and A. E. Gorbalenya.** 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J.Mol.Biol.* **331**:991-1004
75. **Snijder, E. J., M. C. Horzinek, and W. J. Spaan.** 1990. A 3'-coterminal nested set of independently transcribed mRNAs is generated during Berne virus replication. *J.Virol.* **64**:331-338
76. **Snijder, E. J. and J. J. M. Meulenberg.** 1998. The molecular biology of arteriviruses. *J.Gen.Virol.* **79**:961-979
77. **Snijder, E. J., Y. van der Meer, J. Zevenhoven-Dobbe, J. J. Onderwater, J. van der Meulen, H. K. Koerten, and A. M. Mommaas.** 2006. Ultrastructure and origin of membrane vesicles associated with the severe acute respiratory syndrome coronavirus replication complex. *J.Virol.* **80**:5927-5940
78. **Snijder, E. J., A. L. Wassenaar, W. J. Spaan, and A. E. Gorbalenya.** 1995. The arterivirus Nsp2 protease. An unusual cysteine protease with primary structure similarities to both papain-like and chymotrypsin-like proteases. *J.Biol.Chem.* **270**:16671-16676
79. **Snijder, E. J., A. L. Wassenaar, L. C. van Dinten, W. J. Spaan, and A. E. Gorbalenya.** 1996. The arterivirus nsp4 protease is the prototype of a novel group of chymotrypsin-like enzymes, the 3C-like serine proteases. *J.Biol.Chem.* **271**:4864-4871
80. **Steege, D. A.** 2000. Emerging features of mRNA decay in bacteria. *RNA.* **6**:1079-1090
81. **Steffens, S., H. J. Thiel, and S. E. Behrens.** 1999. The RNA-dependent RNA polymerases of different members of the family Flaviviridae exhibit similar properties in vitro. *J.Gen.Virol.* **80** ( Pt 10):2583-2590
82. **Taddeo, B. and B. Roizman.** 2006. The virion host shutoff protein (UL41) of herpes simplex virus 1 is an endoribonuclease with a substrate specificity similar to that of RNase A. *J.Virol.* **80**:9341-9345
83. **Taddeo, B., W. Zhang, and B. Roizman.** 2006. The U(L)41 protein of herpes simplex virus 1 degrades RNA by endonucleolytic cleavage in absence of other cellular or viral proteins. *Proc.Natl. Acad.Sci.U.S.A* **103**:2827-2832
84. **Taddeo, B., W. Zhang, and B. Roizman.** 2009. The virion-packaged endoribonuclease of herpes simplex virus 1 cleaves mRNA in polyribosomes. *Proc.Natl.Acad.Sci.U.S.A* **106**:12139-12144
85. **Te Velthuis, A. J., J. J. Arnold, C. E. Cameron, S. H. van den Worm, and E. J. Snijder.** 2010. The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. *Nucleic Acids Res.* **38**:203-214
86. **Tijms, M. A., L. C. van Dinten, A. E. Gorbalenya, and E. J. Snijder.** 2001. A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus. *Proc.Natl.Acad.Sci.U.S.A* **98**:1889-1894
87. **Ule, J., K. Jensen, A. Mele, and R. B. Darnell.** 2005. CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* **37**:376-386
88. **Urlaub, H., Kl. Hartmuth, S. Kostka, G. Grelle, and R. Luhrmann.** 2000. A general approach for identification of RNA-protein cross-linking sites within native human spliceosomal small nuclear ribonucleoproteins (snRNPs). *J.Biol.Chem.* **275**:41458-41468

89. **Uzan, M.** 2001. Bacteriophage T4 RegB endoribonuclease. *Methods Enzymol.* **342**:467-480
90. **van Aken, D., J. Zevenhoven-Dobbe, A. E. Gorbalenya, and E. J. Snijder.** 2006. Proteolytic maturation of replicase polyprotein pp1a by the nsp4 main proteinase is essential for equine arteritis virus replication and includes internal cleavage of nsp7. *J.Gen.Virol.* **87**:3473-3482
91. **van den Born, E., A. P. Gultyaev, and E. J. Snijder.** 2004. Secondary structure and function of the 5'-proximal region of the equine arteritis virus RNA genome. *RNA.* **10**:424-437
92. **van den Born, E., C. C. Posthuma, A. P. Gultyaev, and E. J. Snijder.** 2005. Discontinuous subgenomic RNA synthesis in arteriviruses is guided by an RNA hairpin structure located in the genomic leader region. *J. Virol.* **79**:6312-6324
93. **van den Born, E., C. C. Posthuma, K. Knoops, and E. J. Snijder.** 2007. An infectious recombinant equine arteritis virus expressing green fluorescent protein from its replicase gene. *J.Gen.Virol.* **88**:1196-1205
94. **van der Meer, Y., H. van Tol, J. K. Locker, and E. J. Snijder.** 1998. ORF1a-encoded replicase subunits are involved in the membrane association of the arterivirus replication complex. *J. Virol.* **72**:6689-6698
95. **van Dinten, L. C., J. A. den Boon, A. L. M. Wassenaar, W. J. M. Spaan, and E. J. Snijder.** 1997. An infectious arterivirus cDNA clone: Identification of a replicase point mutation that abolishes discontinuous mRNA transcription. *Proc.Natl.Acad.Sci.U.S.A* **94**:991-996
96. **van Dinten, L. C., S. Rensen, A. E. Gorbalenya, and E. J. Snijder.** 1999. Proteolytic processing of the open reading frame 1b-encoded part of arterivirus replicase is mediated by nsp4 serine protease and is essential for virus replication. *J.Virol.* **73**:2027-2037
97. **van Dinten, L. C., A. L. M. Wassenaar, A. E. Gorbalenya, W. J. M. Spaan, and E. J. Snijder.** 1996. Processing of the equine arteritis virus replicase ORF1b protein: Identification of cleavage products containing the putative viral polymerase and helicase domains. *J. Virol.* **70**:6625-6633
98. **van Hemert, M. J., A. H. de Wilde, A. E. Gorbalenya, and E. J. Snijder.** 2008. The in vitro RNA synthesizing activity of the isolated arterivirus replication/transcription complex is dependent on a host factor. *J.Biol.Chem.* **283**:16525-16536
99. **van Hemert, M. J., S. H. van den Worm, K. Knoops, A. M. Mommaas, A. E. Gorbalenya, and E. J. Snijder.** 2008. SARS-coronavirus replication/transcription complexes are membrane-protected and need a host factor for activity in vitro. *PLoS.Pathog.* **4**:e1000054
100. **van Marle, G., J. C. Dobbe, A. P. Gultyaev, W. Luytjes, W. J. Spaan, and E. J. Snijder.** 1999. Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. *Proc.Natl.Acad.Sci.U.S.A* **96**:12056-12061
101. **van Marle, G., L. C. van Dinten, W. J. M. Spaan, W. Luytjes, and E. J. Snijder.** 1999. Characterization of an equine arteritis virus replicase mutant defective in subgenomic mRNA synthesis. *J. Virol.* **73**:5274-5281
102. **van Vliet, A. L., S. L. Smits, P. J. Rottier, and R. J. de Groot.** 2002. Discontinuous and non-discontinuous subgenomic RNA transcription in a nidovirus. *EMBO J.* **21**:6571-6580
103. **Vazquez, A. L., J. M. Martin Alonso, R. Casais, J. A. Boga, and F. Parra.** 1998. Expression of enzymatically active rabbit hemorrhagic disease virus RNA-dependent RNA polymerase in *Escherichia coli*. *J.Virol.* **72**:2999-3004
104. **Wassenaar, A. L. M., W. J. M. Spaan, A. E. Gorbalenya, and E. J. Snijder.** 1997. Alternative proteolytic processing of the arterivirus replicase ORF1a polyprotein: Evidence that NSP2 acts as a cofactor for the NSP4 serine protease. *J. Virol.* **71**:9313-9322
105. **Xu, X., Y. Zhai, F. Sun, Z. Lou, D. Su, Y. Xu, R. Zhang, A. Joachimiak, X. C. Zhang, M. Bartlam, and Z. Rao.** 2006. New antiviral target revealed by the hexameric structure of mouse hepatitis virus nonstructural protein nsp15. *J.Virol.* **80**:7909-7917





106. **Yount, B., R. S. Roberts, L. Lindesmith, and R. S. Baric.** 2006. Rewiring the severe acute respiratory syndrome coronavirus (SARS-CoV) transcription circuit: engineering a recombination-resistant genome. *Proc.Natl.Acad.Sci.U.S.A* **103**:12546-12551
107. **Ziebuhr, J. and S. G. Siddell.** 1999. Processing of the human coronavirus 229E replicase polyproteins by the virus-encoded 3C-like proteinase: identification of proteolytic products and cleavage sites common to pp1a and pp1ab. *J.Virol.* **73**:177-185
108. **Ziebuhr, J., E. J. Snijder, and A. E. Gorbalenya.** 2000. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J Gen Virol* **81**:853-879

## SUMMARY

When outside of host cells, viruses are largely biologically inert entities, barely more than pieces of genetic material enclosed in a protein coat and sometimes also wrapped in a lipid envelope. Following infection of a suitable host cell, however, expression of viral genes endows viruses with a remarkable capacity to enlist cellular molecular machineries to the task of producing new viruses. The effects of viral replication on host cell metabolism can be tremendous and often lead to cell death. The flow of the key steps in the replicative cycle of viruses is primarily determined by the viral genome type. The majority of currently known viruses employ RNA genomes, which can be single- or double-stranded, and have messenger RNA polarity (plus-strand) or are complementary to mRNA (minus-strand). Plus-strand RNA virus genomes are readily translated upon their release into the cytoplasm of the infected cell to produce a set of viral proteins. A number of these proteins assemble into membrane-associated RNA-synthesizing complexes that direct viral RNA replication. Key to this process is a virus-encoded RNA-dependent RNA polymerase (RdRp), whose activity is often supported and modulated by virus-encoded protein factors. Together, RdRp and accessory factors are commonly referred to as “the viral replicase”. Replicase subunits of RNA viruses have enzymatic activities and/or nonenzymatic regulatory functions and are frequently multifunctional, a property likely stemming from the limited coding capacity of RNA virus genomes.

One of the most complex RNA virus replicases known to date is encoded by viruses in the order Nidovirales. This order comprises three distantly related families of enveloped plus-strand RNA viruses: *Coronaviridae* (subfamilies *Coronavirinae* and *Torovirinae*), *Arteriviridae*, and *Roniviridae*, which can infect a range of hosts from invertebrates to mammals, including man. The evolutionary relationship among nidoviruses is evident from similarities in gene organization, expression strategies, and the presumed common ancestry of their key replicative enzymes. Nidovirus genome sizes vary between 13-16 kb for *Arteriviridae* and 25-32 kb for *Coronaviridae* and *Roniviridae*, and the latter two families possess the largest RNA genomes currently known. The multidomain replicase is encoded in two large 5'-proximal open reading frames in nidovirus genomes, while the viral structural protein genes are located in the 3'-proximal genomic regions. Translation of the 5'-capped, 3'-polyadenylated nidovirus genomic RNA results in two large replicase polyproteins, pp1a and pp1ab, the latter translated after a -1 ribosomal frameshift. Through processing of the replicase polyproteins by virus-encoded proteases, between 13 and 16 individual viral nonstructural proteins (nsps) are produced in nidovirus-infected cells. In addition to genome replication, the nidovirus replicative machinery also mediates the synthesis of a set of two to ten subgenomic (sg) mRNAs. The production of these molecules secures the expression of the structural protein genes. The sg mRNAs of all nidoviruses are 3'-coterminal with the viral genome, and those synthesized by arteriviruses and coronaviruses also contain a common 5' leader sequence identical to the 5'-proximal region of the genomic RNA. This unique mosaic structure of arterivirus and

coronavirus sg mRNAs is presumed to derive from discontinuous RNA synthesis during the production of the subgenome-length minus-strand templates for sg mRNA synthesis.

The functional characterization of replicase subunits is a prerequisite for illuminating the regulatory mechanisms that govern the diverse activities of nidovirus RNA-synthesizing complexes. Understanding fundamental aspects of viral replication is of critical importance for developing strategies to prevent and treat diseases caused by known and emergent nidoviruses. The work described in this thesis focused on the roles of replicase subunits in coupling different processes in the replicative cycle of equine arteritis virus (EAV). This virus, which is the prototype of the *Arteriviridae*, has the smallest genome among nidoviruses and it is recognized as one of the best studied nidovirus models in terms of molecular biology. The current knowledge on the numerous enzymatic activities found in the nidovirus replicase is largely derived from biochemical and genetic studies of viral nsps and is summarized in Chapter 1. In Chapter 2, a literature review on strategies for spatio-temporal regulation of the eukaryotic plus-strand RNA virus replicative cycle is presented, focusing on the inherent multifunctionality of plus-strand RNA virus genomes as templates for at least three different processes: translation, synthesis of minus-strand RNAs, and production of progeny particles. The implications of emergent common themes for our current understanding of and future research on regulatory events in the nidovirus replicative cycle are summarized.


In the following six Chapters of this thesis, experimental work conducted in this study is presented and discussed. We examined two different EAV replicase subunits: nsp11, containing an enzymatic domain that is conserved across a broad spectrum of nidoviruses (NendoU), and nsp1, a protein specific for arteriviruses having a critical function in sg RNA production. The availability of an EAV full-length cDNA clone greatly facilitates functional studies of the elaborate viral replicative apparatus using reverse and forward genetics. An experimental approach combining site-directed mutagenesis of the virus with in vitro biochemical characterization of a recombinant protein form was employed in Chapters 3 and 4 to gain more insight into the role of nsp11 in the EAV replicative cycle. A nidovirus-specific conserved domain, mapping to the C-terminus of arterivirus nsp11 and coronavirus nsp15, had been identified by comparative sequence analysis. Subsequently, bioinformatics uncovered a distant relationship between this domain and a family of cellular endoribonucleases, suggesting that nidovirus nsps containing this function may be capable of hydrolyzing RNA molecules. The biological importance of this proposed enzymatic activity (now called NendoU) for the replicative cycle of EAV was probed and established in Chapter 3. Deletion of the NendoU-coding sequence from the viral genome rendered EAV RNA accumulation undetectable. Similar phenotypes were obtained upon a number of single amino acid replacements. Substitutions of conserved residues presumed critical for endoribonucleic activity, however, yielded viable mutants, albeit with severe defects in infectious progeny production and small-plaque phenotypes. Upon further analysis of the presumed NendoU active site mutants, evidence for the importance of this domain for optimal EAV RNA synthesis was obtained, and a potential link between NendoU and viral sg mRNA production was identified.

The *in vitro* endoribonuclease activity of the arterivirus NendoU-containing replicase subunit was examined in detail (Chapter 4). In *in vitro* activity assays, bacterially expressed nsp11 from EAV and a distantly related arterivirus, porcine respiratory and reproductive syndrome virus (PRRSV), were found to efficiently cleave RNA 3' of pyrimidines. Both enzymes exhibited a modest preference for cleavage after uridylates, an activity that was greatly reduced by replacement of residues previously implicated in catalysis. Comparative analysis of the distantly related NendoU activity of arterivirus nsp11 and coronavirus nsp15, purified and tested under identical conditions, revealed common and distinct features. Taken together, the data presented in Chapters 3 and 4 reveal that this remarkable RNA-processing activity is common to arteri- and coronaviruses, and provide a solid foundation for future research aimed at understanding the molecular details of NendoU functions during nidovirus replication.

The regulatory roles of EAV nsp1, the only nidovirus protein to date with a demonstrated specific role in sg mRNA production, were explored in Chapters 5 and 6 using reverse and forward genetics. The autoproteolytic release of nsp1 from the replicase polyproteins is mediated by a C-terminal papain-like cysteine proteinase domain (PCP $\beta$ ), and cleavage of the nsp1/2 site was found to be a prerequisite for viral RNA synthesis (Chapter 5). Two more subdomains had previously been identified in nsp1 by comparative sequence analysis - a zinc finger (ZF) domain mapping to the N-terminal portion of nsp1, and an additional cysteine proteinase domain (PCP $\alpha$ ) that is proteolytically inactive. Substitutions of proposed zinc-coordinating residues were found to selectively block EAV sg mRNA accumulation, or to severely decrease the production of infectious progeny without affecting viral RNA accumulation. These data identified nsp1 as a multifunctional replicase subunit with critical functions in at least three consecutive key processes in the EAV replicative cycle: replicase polyprotein processing, sg mRNA production, and virion biogenesis.

In Chapter 6, alanine replacement of non-conserved clusters of polar residues in the nsp1 sequence were tested in the EAV cDNA clone in an attempt to expand the repertoire of viable EAV nsp1 mutants with defects in the protein's regulatory functions. Detailed analysis of mutant and pseudorevertant phenotypes revealed that the relative abundance of EAV mRNAs is tightly controlled by an intricate network of interactions involving all nsp1 subdomains. Furthermore, the quantitative balance among viral mRNA species was found to be critical for efficient production of new virus particles. An assay was developed for the quantitative detection of viral minus-strand RNA species and allowed us to demonstrate that nsp1 likely modulates the accumulation of full-length and subgenome-length minus-strand templates for EAV mRNA synthesis. Collectively, the data presented in Chapters 5 and 6 identify nsp1 as a key coordinator of genome replication, sg mRNA synthesis, and virus production.

To address the molecular details of nsp1 function, the development of *in vitro* assays and solution of the three-dimensional structure of the protein are of significant importance. The different approaches we explored for the production and purification of soluble recombinant nsp1 in *Escherichia coli* are described in Chapter 7. The pitfalls



we encountered upon purification of GST-tagged nsp1 and subsequent tag removal by proteolytic cleavage were largely related to the poor solubility of recombinant nsp1. A successful approach for nsp1 expression from a construct that allowed its autoproteolytic release from a precursor is outlined. This protocol resulted in the purification of soluble, stable recombinant nsp1 that can be used in future research.


In Chapter 8, the implications of our data on EAV nsp1 and nsp11 are discussed in substantial detail in the framework of relevant findings obtained by others. Potential directions for future research are outlined, aimed at identifying the biological substrates and significance of NendoU activity for nidovirus replication and unraveling the molecular mechanisms that underlie the multiple regulatory roles of nsp1.



## SAMENVATTING

Virussen zijn buiten hun gastheercel voornamelijk inerte deeltjes, nauwelijks meer dan fragmenten genetisch materiaal, ingesloten in een eiwitmantel die soms is verpakt in een lipide envelop. Na infectie van een geschikte gastheercel echter verschaft de expressie van hun genoom virussen het opmerkelijke vermogen om de moleculaire machinerie van de cel in the schakelen bij het produceren van nieuwe virussen. Het effect van virale replicatie op het metabolisme van de gastheercel kan enorm zijn en leidt vaak tot de dood van de cel. De volgorde van de voornaamste stappen in de replicatiecyclus van virussen wordt in de eerste plaats bepaald door het type van het virale genoom. De meerderheid van de thans bekende virussen bezit een RNA-genoom, dat enkel- of dubbelstrengs kan zijn en van 'messenger RNA' (mRNA) polariteit (plusstrengs) of complementair daaraan (minstrengs). Plusstrengs RNA-virus genomen worden na hun binnenkomst in het cytoplasma van de geïnfecteerde cel direct vertaald in virale eiwitten. De meeste van deze eiwitten assembleren tot een membraangebonden RNA-synthetiserend enzymcomplex dat de virale replicatie aanstuurt. Sleutelfactor in dit proces is een virus-gecodeerd RNA-afhankelijk RNA-polymerase (RaRp), waarvan de activiteit vaak wordt ondersteund en gereguleerd door virusgecodeerde eiwitfactoren. RaRp en accessoire eiwitten worden gezamenlijk doorgaans "het virale replicase" genoemd. De replicase-onderdelen van RNA-virussen hebben enzymatische activiteiten en/of niet-enzymatische, regulerende taken, en zijn vaak ook multifunctioneel, een eigenschap die waarschijnlijk zijn oorsprong vindt in de beperkte coderende capaciteit van RNA-virusgenomen.

Eén van de meest ingewikkelde, tot op heden bekende, RNA-virus replicases behoort toe aan de virussen in de orde Nidovirales. Deze groep omvat drie ver verwante families van envelop-bevattende, plusstrengs RNA-virussen – *Coronaviridae* (subfamilies *Coronavirinae* en *Torovirinae*), *Arteriviridae*, en *Roniviridae*, die gezamenlijk een scala aan gastheren kunnen infecteren, van ongewervelde dieren tot zoogdieren, inclusief de mens. De evolutionaire relatie tussen nidovirussen uit zich in vergelijkbare genoomorganisaties en genexpressie-strategieën, en de veronderstelde gemeenschappelijke afkomst van hun belangrijkste replicatieve enzymen. De grootte van het nidovirusgenoom varieert van 13 – 16 kb voor *Arteriviridae* tot 25 – 32 kb voor *Coronaviridae* en *Roniviridae*, en daarmee bezitten de laatstgenoemde twee families de grootste RNA-genomen die momenteel bekend zijn. Het multidomein replicase is gecodeerd in de twee grote, 5'-gelegen open leesramen in het nidovirusgenoom, terwijl het 3'-gelegen genoomdeel de genen voor de virale structurele eiwitten bevat. Vertaling van het 5'-cap-dragende en 3'-gepolyadenyleerde nidovirusgenoom resulteert in twee grote replicase polyproteïnen, pp1a en pp1ab, waarvan de laatste vertaald wordt na een -1 ribosomale 'frameshift'. Dankzij klieving van de replicase polyproteïnen door virusgecodeerde proteases worden in nidovirus-geïnfecteerde cellen tussen de 13 en 16 individuele virale niet-structurele eiwitten ('nonstructural proteins'; nsps) geproduceerd. Naast genoomreplicatie verzorgt de replicatiemachinerie van nidovirussen ook de synthese van een set van twee tot tien subgenome (sg) mRNAs. De productie van deze moleculen verzekert de expressie van de



genen die de structurele eiwitten coderen. De sg mRNAs van alle nidovirussen zijn 3'-co-terminaal met het virale genoom, en die van arterivirussen en coronavirussen bevatten aan het 5'-uiteinde tevens een gemeenschappelijke 'leader' sequentie, die identiek is aan het 5'-uiteinde van het genoom-RNA. Aangenomen wordt dat de unieke mozaïekstructuur van arterivirale en coronavirale sg mRNAs voortkomt uit een proces van discontinue RNA-synthese, dat plaatsvindt tijdens de productie van de subgenome min-strengen die gebruikt worden als matrijs voor sg mRNA-synthese.

De functionele karakterisering van replicase-onderdelen is essentieel voor het in kaart brengen van de regulerende mechanismen die de diverse activiteiten van het RNA-synthetiserende enzymcomplex van nidovirussen controleren. Het begrijpen van de fundamentele aspecten van virale replicatie is van doorslaggevend belang voor het ontwikkelen van strategieën ter voorkoming en behandeling van ziekten, veroorzaakt door zowel bekende als onverwacht opduikende nidovirussen. Het werk beschreven in dit proefschrift concentreert zich op de functies die specifieke replicase-onderdelen hebben in het koppelen van verschillende processen tijdens de vermeerderingscyclus van het equine arteritis virus (EAV). Dit virus, het prototype van de *Arteriviridae* familie, heeft het kleinste genoom onder de nidovirussen en wordt gezien als één van de best bestudeerde nidovirus-modellen in termen van moleculaire biologie. De huidige kennis van de talrijke enzymatische activiteiten die aanwezig zijn in het nidovirus-replicase is voornamelijk verkregen door de biochemische en genetische karakterisering van virale nsps en wordt samengevat in hoofdstuk 1. Hoofdstuk 2 bevat een literatuuroverzicht van de strategieën die eukaryote plusstrengs RNA-virussen gebruiken om hun vermenigvuldigingscyclus te reguleren in ruimte en tijd. Er is daarbij speciale aandacht voor de intrinsieke multifunctionaliteit van plusstrengs RNA-virusgenomen als matrijs in tenminste drie verschillende processen: translatie, synthese van minstrengs RNAs, en de productie van nieuwe virusdeeltjes. Het belang van de gemeenschappelijke thema's die zichtbaar worden voor ons huidige begrip van en toekomstig onderzoek aan regulerende stappen in de vermeerderingscyclus van nidovirussen worden samengevat.

In de volgende 6 hoofdstukken van dit proefschrift wordt het experimentele werk, uitgevoerd in dit promotieonderzoek, gepresenteerd en besproken. Twee verschillende EAV replicase-onderdelen werden onderzocht: nsp11, dat een enzymfunctie bevat die geconserveerd is in een breed spectrum van nidovirussen (NendoU), en nsp1, een eiwit specifiek voor arterivirussen met een doorslaggevende functie in sg RNA-productie. De beschikbaarheid van een 'full-length' cDNA kloon voor EAV bevordert functionele studies aan de uitgebreide replicatiemachinerie van dit virus in belangrijke mate, gebruikmakend van zogenaamde 'forward' en 'reverse' genetica. In hoofdstuk 3 en 4 wordt een experimentele aanpak gebruikt die plaatsgerichte mutagenese van het virus combineert met de in vitro biochemische karakterisering van een recombinant eiwit-product om onze kennis te vergroten van de rol van nsp11 in de EAV replicatiecyclus. Een nidovirus-specifiek, geconserveerd domein, gelegen in het C-terminale deel van arterivirus nsp11 en coronavirus nsp15, was reeds eerder geïdentificeerd door vergelijkende sequentieanalyse. Vervolgens onthulde bio-informatica een verre verwantschap

tussen dit domein en een familie van cellulaire endoribonucleases, en suggereerde dat nidovirus nsps die dit domein bevatten in staat zouden kunnen zijn RNA-moleculen te hydrolyseren. Het biologische belang van deze voorspelde enzymatische activiteit (nu NendoU genaamd) voor de vermeerderingscyclus van EAV is onderzocht en bevestigd in hoofdstuk 3. Het verwijderen van het NendoU-coderende gebied uit het virale genoom maakte EAV RNA-accumulatie ondetecteerbaar. Gelijksortige virale phenotypen werden verkregen wanneer een aantal specifieke aminozuren werden vervangen. Mutagenese van geconserveerde aminozuren die verondersteld werden essentieel te zijn voor de endoribonuclease activiteit leverde echter levensvatbare mutanten op, hoewel deze wel ernstige defecten vertoonden op het gebied van de productie van infectieus nageslacht en een zogenaamd 'small-plaque' fenotype hadden. Verdere analyse van deze veronderstelde NendoU 'active site'-mutanten bewees het belang van deze functie voor optimale EAV RNA-synthese en onthulde een potentiële connectie tussen NendoU en de virale sg mRNA-productie.

De in vitro endoribonuclease activiteit van het arterivirus NendoU-bevattende replicase-eiwit werd in detail onderzocht (hoofdstuk 4). Bacterieel geëxprimeerd nsp11 van EAV en een ver verwant arterivirus, het 'porcine respiratory and reproductive syndrome virus' (PRRSV), bleken in in vitro activiteitstesten RNA substraten efficiënt te klieven aan de 3' zijde van pyrimidines. Beide enzymen vertoonden een bescheiden voorkeur voor klieving na uridines, een activiteit die in belangrijke mate kon worden verminderd door mutagenese van aminozuren waarvan al eerder werd aangenomen dat ze van belang zijn voor enzymactiviteit. Vergelijkende analyse van de ver verwante NendoU activiteit van arterivirus nsp11 en coronavirus nsp15, gezuiverd en onderzocht onder dezelfde omstandigheden, toonde zowel gemeenschappelijke eigenschappen als verschillen aan. Gezamenlijk laten de data gepresenteerd in hoofdstuk 3 en 4 zien dat deze opmerkelijke 'RNA-processing' activiteit een gemeenschappelijke eigenschap is van arteri- en coronavirussen, en vormen zij een solide basis voor toekomstig onderzoek gericht op het begrijpen van de moleculaire details van de NendoU functie in de replicatie van nidovirussen.

De regulerende functies van EAV nsp1, het enige tot nu toe bekende nidovirus-eiwit met een aangetoonde specifieke functie in sg mRNA-productie, zijn onderzocht in hoofdstuk 5 en 6, gebruikmakend van 'forward' en 'reverse' genetica. De autoproteolytische klieving van nsp1 van de replicase polyproteïnen wordt uitgevoerd door een C-terminaal papaine-achtig cysteine proteïnase domein (PCP $\beta$ ), en klieving tussen nsp1 en nsp2 is een voorwaarde voor virale RNA-synthese (hoofdstuk 5). Twee andere domeinen van nsp1 zijn eerder geïdentificeerd door vergelijkende sequentie-analyse – een 'zinc finger' (ZF) domein in het N-terminale deel van nsp1, en een tweede cysteine proteïnase domein (PCP $\alpha$ ) dat proteolytisch inactief is. Mutagenese van de veronderstelde zink-coördinerende aminozuren bleek te resulteren in een selectieve blokkade van EAV sg mRNA accumulatie, of in een sterke reductie van de productie van infectieus nageslacht zonder dat de RNA accumulatie was veranderd. Deze data identificeren nsp1 als een multifunctioneel onderdeel van het replicase met kritische functies in tenminste drie op-

eenvolgende sleutelprocessen in de EAV vermeerderingscyclus: replicase polyproteïne klieving, sg mRNA-productie, en virion biogenese.

In hoofdstuk 6 werden EAV mutanten getest waarin niet-geconserveerde groepjes polaire aminozuren in nsp1 waren vervangen door alanine, dit in een poging de collectie levensvatbare nsp1 mutanten met defecten in de regulerende functies van het eiwit uit te breiden. Gedetailleerde analyse van mutanten en pseudorevertanten maakte duidelijk dat de relatieve hoeveelheden van de EAV mRNAs strikt worden gereguleerd door een ingewikkeld netwerk van interacties waarbij alle subdomeinen van nsp1 betrokken zijn. Verder bleek de kwantitatieve balans tussen de virale mRNAs kritisch te zijn voor efficiënte productie van nieuwe virusdeeltjes. Er werd een protocol ontwikkeld voor de kwantitatieve detectie van virale minstrengs RNA-moleculen en dit maakte het mogelijk om aan te tonen dat nsp1 waarschijnlijk de accumulatie moduleert van de 'full-length' en subgenome minstrengs RNAs die als matrijs gebruikt worden om de EAV mRNAs te maken. Gezamenlijk identificeren de data in hoofdstuk 5 en 6 nsp1 als een centrale coördinator van genoomvermeerdering, sg mRNA-synthese en virusproductie.

Om de moleculaire details van de nsp1-functies te kunnen ontrafelen, zullen de ontwikkeling van in vitro test systemen en het bepalen van de drie-dimensionale structuur van het eiwit van groot belang zijn. De verschillende benaderingen die zijn getest voor de productie en zuivering van oplosbaar recombinant nsp1 in *Escherichia coli* zijn beschreven in hoofdstuk 7. Er deden zich problemen voor bij de zuivering van GST-gelabeld nsp1, en de daaropvolgende proteolytische klieving van het GST-gedeelte, die grotendeels te wijten waren aan de slechte oplosbaarheid van recombinant nsp1. Er wordt echter ook een succesvolle aanpak voor nsp1 expressie beschreven, waarbij gebruik werd gemaakt van een construct dat autoproteolytische klieving van nsp1 vanaf een groter expressieproduct mogelijk maakt. Deze aanpak resulteerde in de zuivering van oplosbaar, stabiel recombinant nsp1 dat in toekomstig onderzoek gebruikt kan worden.

In hoofdstuk 8 worden de conclusies van onze analyse van EAV nsp1 en nsp11 in detail besproken tegen de achtergrond van relevante, door anderen verkregen resultaten. Mogelijke richtingen voor toekomstig onderzoek worden besproken, onderzoek gericht op de identificatie van de biologische substraten en de betekenis van NendoU activiteit voor nidovirusrelicatie, en op het ontrafelen van de moleculaire mechanismen achter de diverse regulerende functies van nsp1.

## CURRICULUM VITAE

Danny Nedialkova was born on April 3, 1983 in Sofia, Bulgaria. After obtaining her high school diploma from the First English Language School (Sofia) in 2001, she was admitted to study Biotechnology at the University of Sofia in the same year. In the course of the first semester, she transferred to the Università degli Studi di Perugia (Italy), where she continued her studies in a program offered jointly by twelve universities from nine European countries. As part of this program, she completed research internships at the Intercollegiate Faculty of Biotechnology at the University of Gdansk (Poland) and the Department of Medical Microbiology at the Leiden University Medical Center. In July 2004, she received the European First Level Degree "Job Creation Oriented Biotechnology" *cum laude*. She was a PhD student at the University of Leiden from October 2004 until December 2009. The work described in this thesis was carried out during this period under the supervision of Prof. Dr. E.J. Snijder and Prof. Dr. A.E. Gorbalenya at the Molecular Virology Laboratory of the Department of Medical Microbiology at the Leiden University Medical Center. From July 2010 she will take up a postdoctoral research fellow position at the Max Planck Institute for Molecular Biomedicine in Münster, Germany.

