The molecular biology of arteriviruses

Eric J. Snijder and Janneke J. M. Meulenberg

1 Department of Virology, Leiden University Medical Center, AZL P4-26, PO Box 9600, 2300 RC Leiden, The Netherlands
2 Department of Virology, Institute for Animal Science and Health (ID-DLO), Houribweg 39, 8200 AJ Lelystad, The Netherlands

Introduction

During the Xth International Congress of Virology (Jerusalem, August 1996), the International Committee on the Taxonomy of Viruses (ICTV) formally approved two proposals that marked the end of the taxonomic peregrination of the arteriviruses (Cavanagh et al., 1994; Cavanagh, 1997). First of all, equine arteritis virus (EAV), lactate dehydrogenase-elevating virus (LDV), porcine reproductive and respiratory syndrome virus (PRRSV), and simian haemorrhagic fever virus (SHFV) were united in a newly established family of enveloped positive-stranded RNA viruses, the Arteriviridae. This decision ended their existence as ‘unclassified toga- or flaviviruses’ or members of a ‘floating genus’. Secondly, the novel family Arteriviridae was united with the Coronaviridae (containing the genera Coronavirus and Torovirus) in the newly established order of the Nidovirales.

The establishment of an arterivirus family acknowledges the many unique biological and molecular properties of its four present members. On the other hand, the affiliation with the coronaviruses recognizes an intriguing ancestral relationship at the level of genome organization and replication. The genome sequences of EAV, PRRSV, LDV and SHFV were instrumental in uniting and reclassifying the arteriviruses (den Boon et al., 1991; Godeny et al., 1993; Meulenberg et al., 1993b; Murtaugh et al., 1995; Palmer et al., 1995; Smith et al., 1997). The availability of cloned arterivirus sequences has rapidly increased our understanding of the properties and functions of the viral RNAs and proteins. The recently developed infectious cDNA clones of EAV and PRRSV will allow an even more sophisticated analysis of the arterivirus replication cycle (van Dinten et al., 1997; de Vries et al., 1997; Meulenberg et al., 1998).

After a brief introduction, this review will focus on the molecular biology of the members of the arterivirus family. For detailed information about the clinical, epidemiological and immunological aspects of arterivirus infections the reader is referred to various recent reviews, which also contain references to additional original publications (e.g. Plagemann & Moennig, 1992; Wensvoort et al., 1992; Meredith, 1993; Timoney & McCollum, 1993; de Vries, 1994; Plagemann, 1996).

Arteriviruses and arterivirus disease

EAV-induced disease may have been documented for the first time around the end of the nineteenth century (Pottle, 1888; Clark, 1892). Three of the four currently known arteriviruses (EAV, LDV and SHFV) were first isolated and characterized 30–40 years ago (Doll et al., 1957; Riley et al., 1960; Palmer et al., 1968; Tauraso et al., 1968). The exception is the porcine arterivirus PRRSV, of which different strains emerged (apparently independently) in the USA and Europe about 10–15 years ago (Terpstra et al., 1991; Wensvoort et al., 1991; Collins et al., 1992). The outcome of arterivirus infection can range from an asymptomatic, persistent carrier state to abortion or lethal haemorrhagic fever. For the replication of all arteriviruses, macrophages appear to be the primary target cell (Plagemann, 1996). The host range of the currently known arteriviruses is restricted to horses and donkeys (EAV), pigs (PRRSV), mice (LDV), and several genera of African and Asian monkeys (SHFV). Recently, chickens and mallard ducks which were exposed to PRRSV in drinking water were reported to shed the virus in their faeces, suggesting that they are susceptible to PRRSV infection (Zimmerman et al., 1997). In nature, EAV and PRRSV are assumed to be transmitted primarily via the respiratory route (McCollum et al., 1971; Wensvoort et al., 1992; Meredith, 1993; Timoney & McCollum, 1993). The severity of disease can vary greatly and seems to be influenced by the strain of the virus and the condition and age of the infected animal. EAV and PRRSV infections can be subclinical, but can also lead to a variety of symptoms, including respiratory distress, transient fever and necrosis of small muscular arteries, from which the name of the family prototype, EAV, was derived. Both viruses can cause abortion in pregnant animals, a property which contributed significantly to their initial identification (Doll et al., 1957; Terpstra et al., 1991). Another common trait is their persistent presence in the semen of infected males, which makes sexual transmission an important secondary route of infection (Timoney & McCollum, 1993; Swenson et al., 1994). Recently, PRRSV was reported to replicate in testicular germ cells such
as spermatids and spermatocytes, a property that may be primarily responsible for the excretion of infectious virus in semen (Sur et al., 1997).

SHFV was first isolated in 1964, after devastating outbreaks of haemorrhagic fever in colonies of captive macaque monkeys (Palmer et al., 1968; Tauraso et al., 1968). The virus appears to be endemic among several genera of African monkeys, in which it causes an asymptomatic persistent infection (London, 1977; Gravell et al., 1986b). The epizootics and disease in macaques, during which mortality approached 100%, are probably explained by accidental SHFV transmission from African monkeys to macaques, in combination with an unusual sensitivity of the latter to SHFV infection.

The name of the mouse arterivirus LDV is derived from the increase in the level of lactate dehydrogenase which the virus causes in the blood of infected mice (Riley et al., 1960). LDV has been used extensively as an in vivo research model (Cafruny, 1989; Plagemann & Moennig, 1992). The virus specifically infects a renewable subpopulation of macrophages and is, for unknown reasons, able to escape immune surveillance. Usually, this persistent infection is largely asymptomatic, but in certain inbred mouse strains neurovirulent LDV variants can cause fatal age-dependent polioimmunitis, a complex syndrome which is often linked to the presence of ectropic murine leukemia virus and which leads to the destruction of specific motor neurons by cytoidal LDV replication (Contag & Plagemann, 1989).

**The cell biology of arterivirus infection and replication**

In cell culture, arteriviruses have a very restricted host cell specificity. LDV grows in primary cultures of mouse macrophages, but macrophage and other cell lines fail to support replication (Plagemann & Moennig, 1992). SHFV and PRRSV also replicate in macrophages and, furthermore, in African green monkey kidney cells (MA-104) and derivatives thereof (CL2621 or MARC-145) (Gravell et al., 1986b; Benfield et al., 1992; Kim et al., 1993). In a comparative study on the isolation of American field strains from PRRSV-infected pigs, porcine alveolar lung macrophages (PAM) were found to be more sensitive than CL2621 cells. Surprisingly, certain strains could only be rescued from CL2621 cells and did not replicate in PAM (Bautista et al., 1993). European PRRSV strains have to be adapted to CL2621 by multiple passages before titres of $10^6$–$10^7$ TCID$_{50}$/ml can be obtained. In contrast to the extreme cell type specificities of the other arteriviruses, EAV replicates efficiently in primary cultures of macrophages and kidney cells, as well as in a variety of cell lines such as baby hamster kidney (BHK-21; Hyllseth, 1969), rabbit kidney (RK-13; McCollum, 1970) and African green monkey kidney (VERO; Konishi et al., 1975) cell lines.

The entry pathway of PRRSV into MARC-145 cells was recently investigated using drugs that affect the pH of intracellular membrane compartments (Kreutz & Ackermann, 1996). It was shown that a low pH in these compartments is required for virus entry, which most likely occurs via the standard endocytotic route. Electron microscopy revealed PRRSV particles contained in relatively small vesicles, which appeared to be clathrin-coated. Similar observations were made during the initial stages of LDV infection in macrophages (Kowalchyk & Plagemann, 1985).

Upon transfection of genomic RNA, arteriviruses can replicate in several cell lines that cannot be infected using virus particles (Inada et al., 1993; Meulenberg et al., 1998; L. C. van Dinten & E. Snijder, unpublished data). Furthermore, cell lines which are resistant to LDV can be infected by murine coronavirus/PRRSV pseudovirions (Even & Plagemann, 1995). Certain cell lines have been reported to become LDV-susceptible when they are first infected with murine leukemia virus (Inada & Yamazaki, 1991), although these results have not been reproduced by others (Plagemann, 1996). These data indicate that cell tropism is as least partly determined by the presence of an as yet unidentified receptor on the cell surface. It was suggested that major histocompatibility complex (MHC) class II la antigens might be involved in LDV entry (Inada & Mims, 1985), but this hypothesis was not supported by subsequent experiments. Contradictory results were also reported, showing that trypsin treatment of macrophages destroyed the LDV receptor with a minimal loss of la antigen (Buxton et al., 1988). In the case of PRRSV, infection of MARC-145 cells could be prevented by prior incubation of virus or cells with heparin, and also by treatment of the cells with heparinase (Jusa et al., 1997). Therefore, it was suggested that a heparin-like molecule on the surface of MARC-145 cells serves as an attachment molecule for PRRSV.

Arteriviruses are assembled by budding of preformed nucleocapsids into the lumen of the smooth endoplasmic reticulum (ER) and/or the Golgi region (Fig. 1) and are released by exocytosis (Magnusson et al., 1970; Wood et al., 1970; Stueckemann et al., 1982; Pol & Wagenaar, 1992). The generation of aberrant nucleocapsid structures and/or unusual elongated virions was observed for various arteriviruses (Stueckemann et al., 1982; Pol & Wagenaar, 1992; Wada et al., 1995). A number of arterivirus envelope proteins accumulate in intracellular membranes, but their role in the assembly process has not yet been studied in detail (de Vries et al., 1995b; Mardassi et al., 1996; Y. van der Meer & E. J. Snijder, unpublished data).

In one-step growth experiments, the maximum release of arterivirus particles is between 10–20 h after infection (Stueckemann et al., 1982; van Berlo et al., 1982; Gravell et al., 1986b; Kim et al., 1993). The maximum titres obtained in cell culture are $10^6$–$10^7$ TCID$_{50}$/ml for PRRSV and SHFV, but may exceed $10^8$ TCID$_{50}$/ml for EAV. The infection of macrophages and cell lines is highly cytoidal (Burki, 1965; Tauraso et al., 1968; Onyekaba et al., 1989b; Pol et al., 1991). The cytopathic effect of EAV, SHFV and PRRSV in cell lines is characterized...
Fig. 1. Electron micrographs of arterivirus particles budding from smooth membranes in (or close to) the Golgi region of infected cells. (a) PRRSV-infected porcine alveolar macrophages. Bar, 43 nm. (b) EAV-infected baby hamster kidney cells (BHK-21). Bar, 21 nm.

by rounding of the cells and detachment from the culture plate surface, which allows the titration of these viruses by endpoint dilution or plaque assays (Hyllseth, 1969; van Berlo et al., 1980; Gravell et al., 1986a; Benfield et al., 1992; Kim et al., 1993; van Nieuwstadt et al., 1996). However, in cultures of mouse macrophages the percentage of LDV-permissive cells is too low to detect their destruction. Therefore, the titration of LDV is still performed in vivo in mice (Plagemann et al., 1992).

A typical feature of arterivirus replication is the formation of paired membranes and double-membrane vesicles (DMVs) at 3–6 h post-infection (Fig. 2a, b; Breese & McCollum, 1970; Wood et al., 1970; Stueckemann et al., 1982; Pol & Wagenaar, 1992; Pol et al., 1997). Although the origin and function of these membrane structures are unclear, they do not appear to be involved in virus assembly. Whether they are involved in viral RNA synthesis remains to be studied. The EAV replicase subunits that contain the putative RNA polymerase and helicase functions are membrane-associated and localize to the perinuclear region, which also contains the DMVs (van Dinten et al., 1996). Biochemical and immunofluorescence (Fig. 2c) assays have revealed that several other replicase subunits are anchored to intracellular membranes in EAV-infected cells (van der Meer et al., 1998). Similar conclusions were drawn from an in vitro translation analysis of parts of the LDV replicase (Faaberg & Plagemann, 1996).

**Genome properties and organization**

The arterivirus genome is a single RNA molecule with a length of between 12.7 and 15.7 kilobases (kb). Its positive-stranded nature was first demonstrated for EAV by transfection of purified virion RNA (van der Zeijst & Horzinek, 1975). The 3′ end of the genome is polyadenylated (van Berlo et al., 1982; Sagripanti, 1985; Brinton et al., 1986; Contag et al., 1986). The SHFV genomic 5′ end was found to contain a type I cap structure (Sagripanti et al., 1986) and, furthermore, cap
analogue was reported to be essential for the in vitro generation of infectious PRRSV and EAV RNA from full-length cDNA templates (Meulenberg et al., 1998; L. C. van Dinten & E. J. Snijder, unpublished data). The polycistronic arterivirus genome (Fig. 3a) contains two large replicate open reading frames (ORFs), which are expressed from the genomic RNA, and a set of six to nine downstream genes, which encode mostly structural proteins and are translated from subgenomic mRNAs (see below). The coding regions of the genome are flanked by 5′ and 3′ non-translated regions (NTRs) of 156–221 nucleotides (nt) and 59–117 nt, respectively.

Full-length genomic sequences [sizes excluding poly(A) tails] have been reported for the EAV Bucyrus strain (12704 nt, EMBL database accession no. X53459; den Boon et al., 1991), the Lelystad strain of PRRSV (14104 nt, accession no. M96262; Meulenberg et al., 1993b) and two LDV strains (LDV-C: 14222 nt, accession no. L13298, Godeny et al., 1993; and LDV-P: 14104 nt, accession no. U15146, Palmer et al., 1995). Furthermore, the complete structural protein-coding regions of the genomes of SHFV (accession no. U63121; Smith et al., 1997) and a number of additional PRRSV isolates (Conzelmann et al., 1993; Murtaugh et al., 1995; Kapur et al., 1996) have been published. With the exception of a small region of the SHFV genome, a fairly consistent genome organization was described for all arteriviruses (Fig. 3a). The two replicate ORFs (1a and 1b) occupy approximately three-quarters of the genome. The number of amino acids (aa) encoded by ORF1a is quite variable (between 1727 and 2396). The ORF1b-encoded polypeptide is considerably more conserved, both in size (1410–1463 aa) and in sequence. The size variation of the 5′ half of ORF1a largely explains the differences in genome size between EAV, LDV and PRRSV (Fig. 3a). Downstream of the replicate gene, the EAV/LDV/PRRSV genomes all contain a set of six smaller ORFs, most of which have both 5′- and 3′-terminal sequences that overlap with neighboring genes. The recently published sequence of the 3′-terminal 6 kb of the SHFV genome was surprising in the sense that three additional ORFs (covering about 1.6 kb) were detected in the region between ORF1b and the equivalent of EAV/LDV/PRRSV ORF5 (Fig. 3a; Smith et al., 1997). On the basis of limited sequence similarities, it was proposed that these ORFs have arisen from the duplication of ORFs 2–4 by an RNA recombination event (Godeny et al., 1998). Although it is still unclear whether the additional SHFV ORFs are functional, their presence certainly complicates the establishment of a consistent nomenclature for the arterivirus genes and proteins derived from the 3′ end of the genome. For the purpose of this review, we will use the names membrane (M) protein and nucleocapsid (N) protein for the respective products of ORFs 6 and 7 of EAV/LDV/PRRSV and their equivalents encoded by SHFV ORFs 8 and 9 (Smith et al., 1997). The proteins encoded by EAV/LDV/PRRSV ORFs 2–5 and SHFV ORFs 2a–7 will be referred to as ‘the ORF n protein’.

Recently, the presence of an additional small ORF upstream of EAV ORF1a was noticed (nt 14–124, encoding a hypothetical 37 aa peptide; Kheyar et al., 1996). If translation initiation so close to the 5′ end of the RNA indeed occurs, this ORF could be expressed from all EAV mRNAs, since it is entirely located within the 5′ common leader sequence, which is shared by the genome and all subgenomic mRNAs (see below). Interestingly, the SHFV (Zeng et al., 1995) and PRRSV (Lelystad strain; Meulenberg et al., 1998) genomes also contain potential small ‘intraleader ORFs’ [nt 35–73 (13 aa) and nt 1–38 or 4–38 (13 or 12 aa), respectively]. In the 5′ leader of a number of eukaryotic mRNAs such short ORFs regulate the rate of translation of a downstream reading frame (Kozak, 1991). Although the functionality of the various arterivirus intraleader ORFs remains to be proven, it is interesting to note that similar intraleader ORFs can arise during persistent coronavirus infections (Hofmann et al., 1993; Chen & Baric, 1995).

**Genome translation and replication**

By definition, the genome of a positive-stranded RNA virus fulfills a dual role in both storage and expression of genetic information. Thus, the reproduction of the arterivirus genomic
RNA is a combined process of genome replication and mRNA transcription. For simplicity, we will in this review use the term ‘replication’ for the process of genome (mRNA1) synthesis, whereas the term ‘transcription’ will be used to refer to generation of the subgenomic mRNAs (see below).

The arterivirus replication cycle (Fig. 3b) starts with the expression of the replicase gene from the incoming genome. Replicase ORF1a and ORF1b, which is in the −1 reading frame relative to ORF1a, are both expressed from the genomic mRNA. ORF1b translation requires a ribosomal frameshift just before ORF1a translation is terminated (den Boon et al., 1991). The ORF1a/ORF1b overlap region contains two signals which are assumed to promote this event (Jacks et al., 1988; Brierley, 1995): a so-called ‘slippery’ sequence, which is the actual frameshift site, and a downstream RNA pseudoknot structure (Fig. 4). In EAV, the putative shift site 5′ GUUAAC 3′ is followed immediately by the ORF1a termination codon. In PRRSV and LDV, one additional codon is present between these two elements. The predicted stems 1 and 2 of the RNA pseudoknot typically consist of 11–12 and 6–7 base pairs, respectively (den Boon et al., 1991; Godeny et al., 1993; Meulenberg et al., 1993b). The functionality of the arterivirus frameshift signal has only been demonstrated for EAV: using a reporter gene construct, a frameshift efficiency of 15–20% was observed (den Boon et al., 1991).

Little is known about the arterivirus RNA sequences required for genome replication. An important reason is the fact that, in contrast to, for example, coronaviruses, defective (interfering) genomes have not been described for any of the arteriviruses. In the 3′ coterminal, but they also have a common 5′ leader sequence (with the possible exception of the genus Torovirus in the family Coronaviridae; Snijder & Horzinek, 1993; de Vries et al., 1997). The latter is derived from the 5′ end of the genome and is fused to the subgenomic mRNA ‘bodies’ by a discontinuous transcription mechanism. Previously, it was proposed that the mode of arterivirus subgenomic mRNA transcription might be fundamentally different from that of coronaviruses (van Berlo et al., 1982; de Vries et al., 1990; Chen et al., 1993). However, data have recently accumulated which indicate that the subgenomic mRNA transcription mechanisms of the two groups are essentially similar. Obviously, this idea is also supported by the common ancestry of the arteri- and coronavirus replicase genes (den Boon et al., 1991; Snijder & Spaan, 1995; de Vries et al., 1997). Although comparative studies revealed some differences between the two virus groups, similar UV-inactivation kinetics were observed for coronavirus and arterivirus subgenomic mRNA transcription (den Boon et al., 1995b). Late in infection, the UV target sizes of subgenomic mRNAs were somewhat larger than their physical lengths, but the results ruled out conventional cis-splicing as the major mechanism of subgenomic mRNA production. As in coronaviruses, conserved leader–body junction sequences have been identified in arterivirus mRNAs (see below; de Vries et al., 1990; Chen et al., 1993; Meulenberg et al., 1993a; Zeng et al., 1994).
Fig. 5. (a) Overview of the nested set structure of arterivirus plus- (white) and minus- (black) strand RNAs. Shown are the genomic RNA (RNA1) and the largest (RNA2) and smallest (RNA7) subgenomic RNAs. The locations of the (+) JS sequences (black boxes) and (-) JS sequences (white boxes) are indicated. The leader and anti-leader sequences are represented by hatched boxes. (b, c) Transcription models illustrating the discontinuous step during nidoviral mRNA synthesis, which has been postulated to occur either during plus-strand synthesis (b) or during minus-strand synthesis (c) (reviewed for coronaviruses by van der Most & Spaan, 1995; Sawicki & Sawicki, 1995; Lai & Cavanagh, 1997; Brian & Spaan, 1997). Initially, discontinuous transcription was proposed to occur at the level of plus-strand RNA synthesis, via (+) leader-priming on complementary (-) body JS sequences (b). Following the detection of a nested set of subgenomic RIs, various models incorporating discontinuous minus-strand RNA synthesis were proposed. As illustrated in (c), the (-) body JS sequences may be the signal for a ‘jump’ of the nascent subgenomic minus strand to the 5′ end of the plus template. This would generate a subgenomic minus strand which could function as a template for mRNA synthesis. According to an alternative model (van der Most & Spaan, 1995), the JS sequences in the plus template function may function as minus-strand transcription attenuators, after which leader priming on the 3′-terminal (-) JS sequence of the terminated minus-strand transcript could be used to generate a subgenomic plus-strand RNA.

Fig. 6. Comparison of arterivirus leader–body junction site (JS) sequences. An alignment of the (+) leader JS and the consensus mRNA JS sequences of the four arteriviruses is shown. Conserved nucleotides are boxed. At other positions the most common nucleotide is indicated in upper case, whereas deletions (-) and exceptions (lower case) are indicated below. A small number of mRNA JS sequences, in which mismatches in the presumed base-pairing between the (+) leader JS and the (-) body JS appear to be compensated for by additional base-pairing possibilities, were not included in the comparison. The position of the ORF1a initiation codon downstream of the genomic leader sequence is shown.

1995; den Boon et al., 1996; Godeny et al., 1998). The detection of a set of subgenomic replicative intermediates (RIs) in EAV-infected cells was another important parallel with the coronaviruses (den Boon et al., 1996). These subgenomic RIs contain negative-stranded RNAs that are mirror images of the subgenomic mRNAs (Chen et al., 1994; den Boon et al., 1996) and may therefore be involved in mRNA synthesis (Sethna et al., 1989; Sawicki & Sawicki, 1990; see also below). In this section the positive- or negative-stranded polarity of the various RNA elements will be indicated by (+) and (−), respectively.

As described above and illustrated in Fig. 5(a), the subgenomic mRNAs of arteriviruses and coronaviruses consist of a ‘leader’ and a ‘body’ part which are non-contiguous in the genome sequence: they are transcribed from sequences in the 3′ end and the 5′-terminal one-third of the genomic (−) strand, respectively. The connection between the two segments of the subgenomic mRNA is formed by a conserved junction site (JS) sequence (Fig. 6) which is found both at the 3′ end of the common leader segment and at the 5′ end of the mRNA body. Initially, the observation that the (+) JS at the 3′ end of the leader could base-pair with the (−) JS upstream of each transcription unit in the (−) template led to the proposal of the ‘leader-primed transcription’ model for coronaviruses (Fig. 5b; Baric et al., 1983; Spaan et al., 1983). In this model base-pairing of the (+) and (−) JS is followed by extension of the leader to yield a subgenomic mRNA. However, the detection of a nested set of subgenomic RIs, containing subgenomic (−) RNAs with an anti-leader sequence, indicated that the discontinuous step in subgenomic RNA transcription could occur during (−) strand synthesis instead of (+) strand synthesis (Fig. 5c). A substantial number of coronavirus mRNA transcription models have been put forward (e.g. reviewed by van der Most & Spaan, 1995; Sawicki & Sawicki, 1995; Lai & Cavanagh, 1997; Brian & Spaan, 1997). These are not mutually exclusive and are each supported by a part of the experimental data, most of which were obtained with the mouse hepatitis coronavirus.

Assuming that the role of base-pairing between (+) and (−) JS is the same in the two virus groups, the coronavirus mRNA transcription models can also be applied to arteriviruses. For coronaviruses, it was shown that JS base-pairing is not the only factor determining subgenomic mRNA transcription. Still, the complementarity between the (+) leader JS and the (−) template JS still plays an important role in most transcription models. From the EAV genome sequence it became clear that, in arteriviruses also, the conserved JS sequence [5′ UCAAC 3′ in the case of EAV (+) JS; Fig. 6] cannot be the only factor determining subgenomic mRNA transcription (den Boon et al., 1991). In addition to the (+) leader JS, the genome contains 17 other JS boxes, of which only six appear to function in subgenomic mRNA transcription. A similar situation is found in the other arteriviruses,
although their JS sequences appear to be less conserved than that of EAV (Fig. 6). The mechanism by which the transcriptional complex selects the correct JS sequences remains to be elucidated, but is likely to involve interactions between proteins (viral and/or cellular) and RNA sequences in both nascent strand and template.

An analysis of the mRNA JS sequences of different arteriviruses (Fig. 6) shows that their 3’ side is generally more conserved, and thus probably more important, than their 5’ side (Chen et al., 1993; Meulenberg et al., 1993 a; Godeny et al., 1998). The sequence variation found in the 5’ part of the mRNA JS revealed that the leader-to-body fusion mechanism may be imprecise. A substantial number of arterivirus subgenomic mRNAs consist of subpopulations containing different leader–body junction sequences derived from leader-to-body fusion at the same JS sequence. Analysis of these mRNA JS sequences and the mismatches between the (+) leader JS and the (−) body JS strongly suggested that certain JS nucleotides in the mRNA can be either ‘leader-derived’ or ‘body-derived’. The presumed role of base-pairing in arterivirus leader–body fusion is further supported by a set of EAV mRNA3 variants which are derived from different, closely spaced JS-like sequences: mismatches in the potential base-pairing between the (+) leader JS and the (−) body JS appear to be compensated for by extended base-pairing possibilities between nucleotides from the region immediately 3’ of the leader (+) JS and nucleotides at the 5’ side of the (−) body JS (den Boon et al., 1996). Recently, the use of an alternative JS sequence for leader–body joining was also reported for mRNA7 of the VR2332 isolate of PRRSV (Faaberg et al., 1998). Certain strains of PRRSV and LDV generate an additional mRNA, which is lacking in closely related isolates (Chen et al., 1993; Meng et al., 1996). The biological significance of these additional mRNAs remains to be studied. However, in both these cases the nucleotides at the 5’ side of the (−) body JS also seem to play a role in leader-to-body fusion.

The replicase polyprotein

Translation of the arterivirus replicase gene yields two large polyproteins: the ORF1a protein (187–260 kDa) and the ORF1ab protein (345–421 kDa). The post-translational fate of these precursors has only been studied extensively for EAV and an apparently complete processing scheme has recently been obtained (Fig. 7; Snijder et al., 1994; Wassenaar et al., 1997; van Dinten et al., 1998). The EAV ORF1ab polyprotein is cleaved ten times by three different ORF1a-encoded proteases. In combination with the ribosomal frameshift, this leads to the generation of 12 processing end-products [named nonstructural protein (nsp) 1–12] and a large number of processing intermediates. Many of the latter have considerable half-lives and may therefore have a specific function in the virus replication cycle. The EAV protease domains are located in nsp1, nsp2 and nsp4. These proteases and their corresponding cleavage sites are well-conserved in other arteriviruses (Godyeny et al., 1993; Meulenberg et al., 1993 b; Snijder & Spaan, 1995). Both nsp1 (29 kDa) and nsp2 (61 kDa) contain a cysteine autoprotease domain which mediates, probably exclusively in cis, their rapid release from the polyprotein. The nsp1 protease has been characterized as a papain-like cysteine protease (PCP) (Snijder et al., 1992), with residues Cys-164 and His-230 forming the putative catalytic dyad in EAV. Although the nsp2 cysteine protease (CP; putative active site residues Cys-270 and His-332 in EAV) is most similar to the papain-like group of viral proteases, it possesses a number of unique properties and has been proposed to belong to a new subgroup of viral cysteine proteases (Snijder et al., 1995). Also, the nsp4 serine protease (SP) is a member of a relatively rare group of proteolytic enzymes, the 3C-like serine proteases (Snijder et al., 1996). It combines the catalytic triad of classical chymotrypsin-like proteases (His-1103, Asp-1129 and Ser-1184 in EAV) with the substrate specificity of the 3C-like cysteine proteases, a subgroup of chymotrypsin-like enzymes named after the picornavirus 3C proteases (Gorbunova & Snijder, 1996). Residues in the substrate-binding region of the nsp4 SP are assumed to be responsible for the specificity of the EAV SP for cleavage sites containing a (Glu/Gln)–(Gly/Ser) dipeptide (Snijder et al., 1996; Wassenaar et al., 1997; van Dinten et al., 1998).

Eight SP cleavage sites, which are also conserved in other arteriviruses, have now been identified: five in the C-terminal half of the ORF1a protein and three in the ORF1b-encoded polypeptide. An interesting observation was the fact that, for the processing of the C-terminal part of the ORF1a protein, two alternative pathways can be followed. Either the nsp4/5 (‘major pathway’) or the nsp5/6 and nsp6/7 sites (‘minor pathway’) are processed (Wassenaar et al., 1997). The association of cleaved nsp2 with the nsp3–8 precursor appears to
determine whether the SP can cleave the nsp4/5 site. nsp2 was previously reported to interact very strongly with nsp3 (Snijder et al., 1994). Furthermore, nsp2, nsp3 and nsp5 contain hydrophobic domains which are assumed to play a role in anchoring the EAV replication complex to intracellular membranes (Snijder et al., 1994; van Dinten et al., 1996; van der Meer et al., 1998). Possibly a specific folding or post-translational organization of the nsp2–5 complex is required for proteolysis of the nsp4/5 site. Another interesting consequence of the use of the two pathways is the generation of multiple SP-containing proteins. In a number of processing intermediates (e.g. nsp3–4 and nsp4–5) the SP is linked to a hydrophobic domain and may therefore be membrane-associated. The fully cleaved nsp4 is probably cytoplasmic and was shown to be efficient in processing other parts of the replicase polyprotein in trans (van Dinten et al., 1998).

In addition to the three proteases described for EAV, both LDV and PRRSV possess a fourth nonstructural proteolytic activity (Godeny et al., 1993; Meulenberg et al., 1993b; den Boon et al., 1995a), which is located at the extreme N terminus of the ORF1a protein and has been named PCPα (Fig. 7). Like the downstream PCPβ domain, which is the equivalent of the EAV nsp1 PCP, the LDV/PRRSV PCPα protease is a papain-like cysteine protease. Its predicted active site residues are Cys-76 and His-146/147 and it mediates the rapid liberation of an additional 20/22 kDa cleavage product, named nsp1α (den Boon et al., 1995a). EAV nsp1 (260 aa) is considerably smaller than the corresponding nsp1α–nsp1β region in LDV and PRRSV (384 and 380 residues, respectively). Remarkably, remnants of an inactivated PCPα domain could be identified in the EAV nsp1 sequence. Although a candidate catalytic Cys residue is lacking, explaining the inactivation of the EAV PCPα domain, the sequences around His-122 display convincing sequence similarity with the region surrounding the PCPα active site His-146/147 of LDV/PRRSV. The consequences of the inactivation of the EAV PCPα domain are unclear, but the partial conservation of the inactivated PCPα indicates that this part of the replicase probably contains an additional, non-proteolytic function (den Boon et al., 1995a; Gorbalenya & Snijder, 1996).

With one exception, predictions relating to the functions of the ORF1b-encoded replicase subunits are derived from comparative sequence analysis. Putative RNA-dependent RNA polymerase and nucleoside triphosphate-binding/RNA helicase motifs were identified in nsp9 and nsp10, respectively (Gorbalenya et al., 1989; den Boon et al., 1991; Godeny et al., 1993; Meulenberg et al., 1993b; van Dinten et al., 1996). Although common to positive-stranded RNA viruses, both the level of sequence similarity and the relative position of these domains clearly links the arterivirus replicase to that of coronaviruses (see also below). Furthermore, two additional domains were identified that appear to be unique to nidoviruses: a conserved domain in nsp11 and a Cys/His-rich domain in the N-terminal region of nsp10. The latter was previously proposed to have metal-binding properties (Gorbalenya et al., 1989; Lee et al., 1991) and was recently implicated in a remarkable mRNA transcription defect. The serendipitous replacement of Ser-2429 by Pro, which is located at the C-terminal border of the nsp10 Cys/His-rich domain, rendered the EAV full-length cDNA clone non-infectious by abolishing subgenomic mRNA transcription (van Dinten et al., 1997). However, the genomic RNA encoding this mutant replicase was replicated efficiently, showing that nidovirus replication and subgenomic mRNA transcription are (at least in part) separate processes.

The arterivirus particle

Arteriviruses are spherical, enveloped viruses (Fig. 8a) with a diameter of 40–60 nm (Table 1; Maes et al., 1970; Horzinek et al., 1971; Brinton Darnell & Plagemann, 1975; Trousdale et al., 1975; Wensvoort et al., 1992; Dea et al., 1995). They consist of an isometric core particle (25–35 nm in diameter),

![Arterivirus particle](image)

**Fig. 8.** The arterivirus particle. (a) Electron micrograph of extracellular PRRSV particles. Bar, 23 nm. (b) Schematic representation of the PRRSV particle and the location of the structural proteins GP₁–GP₅, M and N (encoded by ORFs 2–7, respectively). It should be noted that the equivalents of GP₁ and GP₂ have not (yet) been identified as structural components of any of the other arteriviruses.
Table 1. Molecular properties of arterivirus particles

<table>
<thead>
<tr>
<th>Virus</th>
<th>Diameter (nm)</th>
<th>Density (g/cm³ in sucrose)</th>
<th>Genome size (kb)</th>
<th>Known structural proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAV</td>
<td>50–60</td>
<td>1.17</td>
<td>12.7</td>
<td>Gₚ, ORF2 227 +</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gₚ, ORF5 255 +</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M ORF6 162 –</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N ORF7 110 –</td>
</tr>
<tr>
<td>LDV</td>
<td>55</td>
<td>1.13</td>
<td>14.2</td>
<td>VP₃ ORF5 215 +</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M ORF6 172 –</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N ORF7 116 –</td>
</tr>
<tr>
<td>PRRSV</td>
<td>45–55</td>
<td>1.14</td>
<td>15.1</td>
<td>GP₂ ORF2 249 +</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GP₃ ORF3 265 +</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GP₄ ORF4 183 +</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GP₅ ORF5 201 +</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M ORF6 173 –</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N ORF7 128 –</td>
</tr>
<tr>
<td>SHFV</td>
<td>45–50</td>
<td>1.15</td>
<td>15.7</td>
<td>p42 ? 38–46 kDa +</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p54 ? 48–60 kDa +</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M ORF8 162 –</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N ORF9 111 –</td>
</tr>
</tbody>
</table>

which is surrounded by a rather smooth envelope (Fig. 8a). Large spikes, typical for coronavirus, are absent from the arterivirus surface. A brief incubation with non-ionic detergents releases the viral nucleocapsid from its envelope (Horzinek et al., 1984; Meulenberg et al., 1995b). The buoyant density of arteriviruses is 1.13–1.17 g/cm³ in sucrose, whereas sedimentation coefficients of between 214S and 230S have been reported (Hyllseth, 1970; Brinton Darnell & Plagemann, 1975; Trousdale et al., 1975; van der Zeijst & Horzinek, 1975; Sagripanti, 1984; Meulenberg et al., 1995b). The buoyant density of arteriviruses is 1.13–1.17 g/cm³ in sucrose, whereas sedimentation coefficients of between 214S and 230S have been reported (Hyllseth, 1970; Brinton Darnell & Plagemann, 1975; Trousdale et al., 1975; van der Zeijst & Horzinek, 1975; Sagripanti, 1984; Meulenberg et al., 1995b). Arteriviruses can be stored successfully at −70°C or −20°C, but lose their infectivity slowly when stored at 4°C. The half-life of the Lelystad isolate of PRRSV at pH 7.5 was shown to decrease from 140 h at 4°C to 3 h at 37°C (Bloemraad et al., 1994). The virus was stable at pH 6.0–7.5, but a pH increase or decrease resulted in loss of infectivity. Likewise, the stability of EAV and SHFV is influenced by temperature and pH (Burki, 1966; Tauraso et al., 1968; Harry & McCollum, 1981).

A schematic representation of the arterivirus architecture is shown in Fig. 8(b). The RNA genome (12.7–15.1 kb) is encapsidated by a single nucleocapsid protein, yielding a putatively icosahedral core structure. The lipid bilayer which surrounds the nucleocapsid contains three to five additional structural proteins. For all four arteriviruses, three major structural protein components have been identified in purified virions (Brinton Darnell & Plagemann, 1975; de Vries et al., 1992; Faaberg et al., 1995a; Godeny et al., 1995; Meulenberg et al., 1995b; Mardassi et al., 1996): the nucleocapsid protein N (12–15 kDa), the non-glycosylated membrane protein M (16–20 kDa) and the major glycoprotein (24–44 kDa; GP₃ in Fig. 8). The M and N proteins are encoded by the penultimate and ultimate 3' ORFs, respectively (ORFs 6 and 7 of EAV/LDV/PRRSV; ORFs 8 and 9 of SHFV). The major glycoprotein is the N-glycosylated product of ORF5, which was previously designated G₃ in the case of EAV (de Vries et al., 1992), VP₃ for LDV (Brinton Darnell & Plagemann, 1975; Faaberg & Plagemann, 1995) and GP₄ or E for PRRSV (Meulenberg et al., 1995b; Mardassi et al., 1996). The equivalent protein of SHFV is either 42 or 54 kDa, but the corresponding gene has not yet been assigned (Godeny et al., 1995; Smith et al., 1997). The glycoproteins encoded by ORF2 of EAV (G₆; de Vries et al., 1995b) and ORFs 2–4 of PRRSV (GP₂, GP₃ and GP₄; Meulenberg & Petersen-den Besten, 1996; van Nieuwstadt et al., 1996) were shown to be minor structural proteins.

The nucleocapsid protein

The arterivirus N protein is a small basic protein with a molecular mass of 12–15 kDa. It is expressed abundantly in infected cells and constitutes about 20–40% of the protein content of the virion (de Vries et al., 1992; Faaberg & Plagemann, 1995; Godeny et al., 1995; Bautista et al., 1996). The EAV N protein was shown to be phosphorylated (Zeegers et al., 1976). Immunoprecipitation of the N protein from infected cells and constitutes about 20–40% of the protein content of the virion (de Vries et al., 1992; Faaberg & Plagemann, 1995; Godeny et al., 1995; Bautista et al., 1996). The EAV N protein was shown to be phosphorylated (Zeegers et al., 1976). Immunoprecipitation of the N protein from infected cells and constitutes about 20–40% of the protein content of the virion (de Vries et al., 1992; Faaberg & Plagemann, 1995; Godeny et al., 1995; Bautista et al., 1996).
monomers were observed in EAV and LDV particles (de Vries et al., 1995a; Faaberg et al., 1995a). The regions of the N protein which are important for RNA binding or interactions with envelope proteins have not yet been defined.

Sera from (persistently) LDV-infected mice contain a low level of antibodies directed against the N protein (Coutelier et al., 1986; Faaberg & Plagemann, 1995). In contrast, convalescent sera from EAV- and PRRSV-infected pigs contain high titres of anti-N protein antibodies. Therefore, this polypeptide appears to be a suitable antigen for diagnostic tests in the case of EAV (Chirnside et al., 1995; Kheyar et al., 1997) and PRRSV (Meulenberg et al., 1995a; Denac et al., 1997; Rodriguez et al., 1997). Several groups have generated monoclonal antibodies (MAbs) specific for the PRRSV N protein (Nelson et al., 1993; Drew et al., 1995; van Nieuwstadt et al., 1996; Wieczorek-Krohmer et al., 1996; Rodriguez et al., 1997). Some MAbs recognize epitopes that are specific for European or North American isolates, whereas others recognize epitopes conserved in both PRRSV subgroups. One of the latter was mapped to the central region (aa 50–66) of the N protein (Rodriguez et al., 1997).

### The major envelope proteins

The non-glycosylated M protein is the most conserved structural protein of arterviruses. Its hydrophobic profile resembles that of the coronavirus M protein, suggesting a similar topology and membrane-associated function (de Vries et al., 1992; Kuo et al., 1992; Meulenberg et al., 1993b). The N-terminal half of the protein contains three potential membrane-spanning regions which may infer an NexCy orientation. A short stretch of only 10–18 aa is thought to be exposed at the virion surface, which probably explains the poor immunogenicity of this membrane polypeptide. The M protein is assumed to play an important role in virus assembly and budding. It accumulates in the ER, where it forms disulfide-linked heterodimers with the major (ORF5) glycoprotein (de Vries et al., 1995a; Faaberg et al., 1995a; Mardassi et al., 1996; Y. van der Meer & E. J. Snijder, unpublished data). In cells infected with PRRSV or EAV, disulfide-linked M protein homodimers were also observed but these were not incorporated into virions. The Cys residue in the short N-terminal ectodomain of the M protein is most likely involved in the formation of an intermolecular disulfide bridge with a Cys residue in the ectodomain of the ORF5 glycoprotein. Treatment of LDV virions with 5–10 mM DTT, to disrupt disulfide bonds, reduced their infectivity (Faaberg et al., 1995a), suggesting that heterodimers of the M protein and the ORF5 glycoprotein are important for receptor binding.

The membrane topology of the EAV/LDV/PRRSV ORF5 glycoprotein is still unknown, but its hydrophobicity profile is very conserved. The internal hydrophobic domain might span the viral envelope either one or three times. Compared to EAV, the putative ectodomains of the PRRSV and LDV ORF5 proteins are much smaller, but in all three viruses this region contains N-glycosylation sites. In the case of EAV, and probably also LDV, glycosylation of the ORF5 protein occurs by the addition of variable numbers of lactosamine repeats (de Vries et al., 1992; Li et al., 1998).

Neutralizing antibodies in polyclonal sera from LDV-infected mice (Cafruny et al., 1986) or EAV-infected horses (Chirnside et al., 1995a, 1995b) are predominantly directed against the ORF5 glycoprotein. Furthermore, all published LDV- and EAV-neutralizing MAbs are ORF5 protein-specific (Coutelier & Van Snick, 1988; Harty & Plagemann, 1988; Balasuriya et al., 1993; Deregt et al., 1994; Glaser et al., 1995), and MAbs recognizing the PRRSV ORF5 protein were recently reported to neutralize virus infectivity (Pirzadeh & Dea, 1997). EAV-neutralizing horse antibodies and MAbs bind to the putative ectodomain (aa 19–115) of the ORF5 protein (Chirnside et al., 1995a; Balasuriya et al., 1997). Neutralization escape mutants of EAV contained mutations or even substantial deletions in this region of the ORF5 protein (Balasuriya et al., 1995a; Glaser et al., 1995; Balasuriya et al., 1997). In the latter case, the removal of 47 (aa 66–112) or 40 (aa 62–101) residues from the ectodomain did not affect the ability of these escape mutants to infect cells. Recently, the major neutralization epitope of LDV was also mapped to the putative ectodomain (30 aa) of the ORF5 glycoprotein (Li et al., 1998).

As in the case of EAV (Glaser et al., 1995), the number of glycosylation sites in this region differs between strains of LDV (Faaberg et al., 1995b), a property that was recently implicated in LDV persistence. It was postulated that neutralizing antibodies can bind less efficiently to strains containing an extensively glycosylated ORF5 protein (Li et al., 1998), which could explain the ability of these LDV variants to establish persistent infections.

The ORF5 protein of a European PRRSV isolate was reported to induce apoptosis when expressed using a vaccinia virus recombinant (Suarez et al., 1996a). Interestingly, apoptosis was also detected in PRRSV-infected PAM and MA-104 cells and could not be inhibited by Bcl-2. This suggests either the existence of a target which is downstream of Bcl-2 family members, or the use of an alternative and unknown apoptotic pathway (Teodoro & Branton, 1997). At present, the relevance of PRRSV ORF5 protein-induced apoptosis for infection and pathogenesis in vivo is unclear.

### Minor structural glycoproteins

Besides the N and M proteins and the major glycoprotein, which together represent approximately 90–95% of the structural protein content of arterivirus particles, one or more minor envelope proteins have been identified. The EAV and PRRSV ORF2 proteins are present in virions (de Vries et al., 1992; Meulenberg & Petersen-den Besten, 1996) and were designated Gα and GPα, respectively. The ORF2 protein appears to be indispensable for virus replication, since its
deletion rendered the EAV full-length cDNA clone non-infectious (van Dinten \textit{et al.}, 1997). The ORF2 proteins are class I integral membrane proteins that contain complex type N-glycans when incorporated into virions. The PRRSV ORF2 protein in virus particles is a monomer. In infected cells, the EAV ORF2 protein exists as four monomeric species due to the formation of alternative intrachain disulfide bonds between the three luminal cysteines. However, only disulfide-linked homodimers are incorporated into virions. The ORF2 proteins of PRRSV and EAV are retained in the ER upon individual expression \textit{in vivo} (de Vries \textit{et al.}, 1995\textit{b}; Meulenberg \& Petersen-den Besten, 1996), which indicates that they probably need other structural proteins and/or virus assembly for transport from the ER to the Golgi complex.

The proteins encoded by PRRSV ORFs 3 and 4 have been identified as minor glycoprotein constituents of the virus particle (van Nieuwstadt \textit{et al.}, 1996). The corresponding ORFs of EAV and LDV probably also encode glycoproteins, but these have not yet been detected in infected cells or virions. By \textit{in vitro} transcription and translation in the presence and absence of microsomal membranes, it was shown that the LDV ORF3 and ORF4 products are soluble and membrane-bound, respectively (Faaberg \& Plagemann, 1997). Interestingly, MAbs specific for the PRRSV ORF4 protein are neutralizing (van Nieuwstadt \textit{et al.}, 1996). Using chimeric constructs and pepscan analysis, the neutralizing MAb-binding domain was mapped to the region between aa 39 and 79 (Meulenberg \textit{et al.}, 1997). This region is highly heterogeneous among different strains of PRRSV, indicating that it is subject to immunoselection. The biological functions of the ORF4-encoded protein in virus assembly and host cell attachment remain to be elucidated.

**Sequence variation in the structural proteins**

Comparison of the arterivirus structural protein sequences shows that the M and N proteins are more conserved than the glycoproteins. Of the latter, the EAV/LDV/PRRSV ORF5 glycoprotein appears to be most conserved. These sequence comparisons also revealed that PRRSV, LDV and SHFV are more closely related to each other than they are to EAV. Whereas only one SHFV strain (Smith \textit{et al.}, 1997) and two LDV strains (Godeny \textit{et al.}, 1993; Palmer \textit{et al.}, 1995) have been sequenced, more extensive surveys have been performed for a variety of EAV and PRRSV field isolates. Sequence comparisons revealed that, among EAV field isolates also, the M and N protein sequences (Chirnside \textit{et al.}, 1994) are more conserved than those of the ORF5 and ORF2 glycoproteins (Balasuriya \textit{et al.}, 1995\textit{b}; Glaser \textit{et al.}, 1995; Hedges \textit{et al.}, 1996; Lepage \textit{et al.}, 1996). Furthermore, the ORF5 protein sequence of these field isolates was more variable (89.8–99.6\% identity) than the ORF2 protein sequence (91.6–100\% identity). A comparative study of different European and North American PRRSV isolates showed that the ORF5 glycoprotein is also the most heterogeneous structural polypeptide in this virus (Mardassi \textit{et al.}, 1995; Kapur \textit{et al.}, 1996; Suarez \textit{et al.}, 1996\textit{b}). Interestingly, the ORF5 glycoproteins of different European and US strains of PRRSV are about as closely related to each other (52–55\% identity) as they are to the corresponding LDV protein (47\% identity). Since the ORF 2–7 protein sequences of European and North American PRRSV isolates differ by 21–45\%, the isolates from these two continents are genetically distinct and must have diverged from a common ancestor a relatively long time ago (Murtaugh \textit{et al.}, 1995, 1998).

**Immunology and vaccines**

Arteriviruses can cause a persistent infection, which may last from 2–3 months in PRRSV-infected pigs (Zimmerman \textit{et al.}, 1992) to a lifetime in LDV-infected mice (Onyekaba \textit{et al.}, 1989\textit{a}), EAV-infected stallions (Timoney \& McCollum, 1993) and SHFV-infected patas monkeys (Gravell \textit{et al.}, 1986\textit{b}). The mechanism by which arteriviruses can evade the host immune response is still unknown. Antibodies are detected within a week after infection (Coutelier \textit{et al.}, 1986; Gravell \textit{et al.}, 1986\textit{b}; McCollum, 1986; Nelson \textit{et al.}, 1994). Antibodies raised against EAV are highly neutralizing, their appearance coincides with virus clearance, and they may persist for years (McCollum, 1970; Fukunaga \textit{et al.}, 1981). Therefore, the humoral immune response against EAV is thought to play an important role in protection. On the other hand, low levels of neutralizing antibodies against LDV and PRRSV are detected 1–2 months after infection and do not seem to reduce viraemia (Cafruny \textit{et al.}, 1986; Loemba \textit{et al.}, 1996; Yoon \textit{et al.}, 1996). Virus–antibody complexes formed in LDV-infected mice and PRRSV-infected pigs may even enhance infection, since it was shown \textit{in vitro} that virus uptake by macrophages was increased by the presence of virus-specific antibodies (Cafruny \& Plagemann, 1982; Yoon \textit{et al.}, 1996). Some evidence for a role of maternal antibodies in the protection of off-spring was obtained for LDV and PRRSV (Broen \& Cafruny, 1993; Morrison \textit{et al.}, 1997). The anti-LDV immune response prevents the onset of age-dependent poliomyelitis in mice: it specifically protects motor neurons from LDV infection, without affecting virus replication in permissive macrophages (Harty \textit{et al.}, 1987; Harty \& Plagemann, 1990; Plagemann \& Moennig, 1992). The persistent infection of mice by LDV is thought to be maintained by a slow but continuous regeneration of LDV-permissive macrophages that become infected as soon as they arise (Onyekaba \textit{et al.}, 1989\textit{a}; Rowland \textit{et al.}, 1994). In addition, virus variants that are resistant to neutralization may emerge during chronic infection, and may contribute to the development of a persistent infection (Monteyne \& Coutelier, 1994; Chen \textit{et al.}, 1997). Besides infectious virion–antibody complexes, immune complexes of 150–300S are formed in LDV-infected mice, most likely as a...
consequence of the polyclonal activation of B cells (Cafruny et al., 1986). This activation was not observed in mice immunized with inactivated LDV, suggesting that an active infection of macrophages is required. Although the mechanism of polyclonal B cell activation by LDV is not completely understood, it was proposed that one of the LDV proteins could function as a superantigen that activates helper T cells, or that activation might be induced by the numerous lymphokines secreted by LDV-infected macrophages (Plagemann & Moennig, 1992).

The humoral response against SHFV varies with the strain of monkey and the virus isolate that is tested. The rapid death of macaques after SHFV infection precludes an effective host immune response, but the response in patas monkeys is different. SHFV strains that cause an acute infection in patas monkeys induce a high titre of neutralizing antibodies (Gravell et al., 1986), which correlates with the complete clearance of the virus from the circulation. On the other hand, SHFV strains that persistently infect these monkeys induce only low titres of non-neutralizing antibodies. Interestingly, patas monkeys persistently infected with SHFV can completely clear the virus from the circulation. The neutralizing antibodies raised against the superinfecting strain do not cross-neutralize the persistently infecting SHFV strain (Gravell et al., 1986b). These findings suggest that, besides the humoral immune response, cell-mediated immunity also plays a role in elimination of SHFV.

Although the cellular immune responses against EAV and SHFV have not yet been studied in detail, initial experiments with LDV indicate that cytotoxic T lymphocytes (CTLs) that specifically lyse infected macrophages are elicited in infected mice (Even et al., 1995). However, these CTLs disappeared by 30 days post-infection and did not suppress LDV replication, since progeny virus was already produced and secreted before the infected macrophages were lysed by the CTLs. Other investigators detected cytotoxic and helper T cell responses for up to at least 250 days post-infection, but again no reduction in LDV levels in the blood was observed (van den Broek et al., 1997). A T cell response to PRRSV in pigs was first detected at 4 weeks post-infection, peaked 3 weeks later, and declined at 11 weeks post-infection (Bautista & Molitor, 1997). Therefore, the role of cell-mediated immunity in the protection against arterivirus infections remains to be established.

For both PRRSV and EAV, live attenuated (RespPRRS, Boehringer; Arvac, Ford Dodge Laboratories) and killed vaccines (Cyblue, Labouritarios Sobrino; Artervac, Ford Dodge Laboratories) are commercially available to protect pigs and horses, respectively. In comparison to the killed vaccines, the attenuated vaccines induce an immunity against disease which lasts longer and is more efficient (McCullum, 1986; Fukunaga et al., 1990; Gorczyca et al., 1995; Nielsen et al., 1997). However, these live vaccines can still be improved, since they do not prevent reinfection and do not allow serological discrimination between vaccinated animals and animals infected with field virus.

**Nidovirus evolution: the common ancestry of the replicase genes of arteri- and coronaviruses**

Genome replication and mRNA transcription are fundamental processes in the virus replication cycle. The sequences of replicase proteins involved in viral RNA synthesis are conserved among seemingly disparate groups of RNA viruses. In nidoviruses, related replicase genes and replication strategies have been coupled to seemingly unrelated sets of structural genes (Snijder & Horzinek, 1993; Snijder & Spaan, 1995; de Vries et al., 1997). The nidovirus replicase is large in the case of arterviruses (345–420 kDa; den Boon et al., 1991; Godeny et al., 1993; Meulenberg et al., 1993b), and extremely large in the case of coronaviruses (740–810 kDa; Boursnell et al., 1987; Gorbalevtsa et al., 1989; Lee et al., 1991; Herold et al., 1993; Bonilla et al., 1994; Eleouet et al., 1995) and probably also toroviruses (Snijder & Horzinek, 1993). The arrangement of conserved replicase domains is unique to the nidoviruses, and clearly separates them from other groups of plus-stranded RNA viruses. In the most conserved replicase domains, up to 30% amino acid sequence identity is found between arterviruses and coronaviruses, a percentage that cannot be due to convergent evolution. Also, the arrangement of conserved domains within the replicase polyprotein is unique, for example only in the nidovirus replicase is the helicase domain located downstream of the polymerase motif. The polymerase motif also carries another nidovirus trademark: the substitution of the canonical GDD by SDD.

All nidovirus replicases are processed by multiple ORF1a-encoded proteases. Despite the fact that our understanding of the processing of the much larger coronavirus replicase is far from complete, a common pattern for coronaviruses and arterviruses is emerging (Snijder & Spaan, 1995; de Vries et al., 1997). In comparable positions, the ORF1a proteins of both virus groups contain a domain belonging to the protease superfamily that comprises the chymotrypsin-like and picornavirus 3C-like proteolytic enzymes (Gorbalenya et al., 1989; Liu & Brown, 1995; Lu et al., 1995; Ziebuhr et al., 1995; Snijder et al., 1996). The presence of hydrophobic domains on either side of this ‘main protease’ is another property shared by arterviruses and coronaviruses. The functions encoded by the central region of ORF1a to the 3’ end of ORF1b appear to constitute the ‘core replicase’: the conserved domains (main protease–polymerase–metal-binding domain–helicase–conserved C-terminal ORF1b domain) are within this region (Fig. 7). Only small insertions and deletions are found in this part of the replicase, which is processed by the chymotrypsin/3C-like protease. The N-terminal half of the ORF1a protein, on the other hand, is quite variable and this property is one of the main reasons for the genome size variation observed among corona- and arterviruses. In this region, both corona- and arterviruses contain protease domains which have been shown to belong to the papain-like cysteine protease superfamily.
The variability of the N-terminal half of the ORF1a protein suggests that species-specific, rather than group-specific, functions have been added to the set of basic functions provided by the ‘core replicase’.

The coupling of different gene sets to the same replicase module has been explained by the recombination of complete genes or gene sets (Zimmer, 1987; Goldbach & Wellink, 1988; Strauss & Strauss, 1988). The RNA recombination frequency during coronavirus replication is remarkably high (see reviews by Spaan et al., 1988; Lai, 1996), a characteristic that may be determined by replicase properties. The different sets of structural genes which have been linked to the nidovirus ‘core replicase module’ indicate that this characteristic may be shared by other members of the newly established order, supporting this concept of ‘modular evolution’. Direct evidence for multiple recombination events during torovirus evolution has already been obtained (Snijder et al., 1991; Cornelissen et al., 1997). The presumed duplication of the arterivirus nsp1 protease domain (den Boon et al., 1995a) and of three ORFs in the SHFV genome (Godyen et al., 1998; Fig. 3a) suggest the occurrence of RNA recombination in arteriviruses also. Together with divergent evolution, a high recombination frequency can account for the diverse composition of nidovirus genomes. It has been hypothesized that the acquisition of a novel nucleocapsid protein gene, giving rise to a virus core with different structural properties, may have created the possibility for an ancestor of the coronavirus/torovirus branch to increase its genome size and diverge from the arterivirus branch (Godyen et al., 1993).

Concluding remarks

In molecular terms, the arteriviruses can be considered viruses with ‘a split personality’. The nidovirus replicase module has been linked to a unique set of structural proteins which assemble into an RNA virus particle that possesses a number of distinctive properties. Thus far, the characteristics of the very small arterivirus N protein (110–128 aa), the only protein component of the viral core, have not been studied in any detail. Likewise, the role of the various envelope proteins is poorly understood. Some of them (the M protein and the major glycoprotein) are important structural components, whereas only small quantities of a number of other envelope proteins are present in arterivirus particles. Understanding the properties of the structural proteins and their role in virus assembly will be one of the major challenges for future molecular biological research on arteriviruses. For these kinds of studies, the recently developed infectious cDNA clones of EAV and PRRSV will be extremely valuable tools. Furthermore, the ability to genetically modify these two economically relevant arteriviruses could lead to the rapid development of (marker) vaccines and arterivirus-based expression vectors. At the level of RNA transcription and replicase function, future research may shed light on the RNA and protein sequences involved in genome replication and encapsidation, the regulation of replicase function by proteolytic processing, the interactions of the arterivirus replication machinery with host cell components, and the mechanism of discontinuous subgenomic mRNA transcription. Thus, the future functional dissection of arterivirus RNA and protein sequences, and comparison with the distantly related coronaviruses, is also likely to increase our understanding of some of the common principles underlying nidovirus replication.

We would like to acknowledge the input of many arterivirus colleagues from the Institutes at Leiden, Lelystad and Utrecht, who have worked with us over the years. In particular we would like to thank Johan den Boon, Judy Bos, Leonie van Dinten, Sasha Gorbalenya, Rob Moormann, Ton van Nieuwstadt, Jan Pol, Willy Spaan, Twan de Vries, Fred Wassenaar and Gert Wensvoort. We are grateful to Margo Brinton, Elmer Godeny, Kay Faaberg, Michael Murtaugh, Peter Plagemann and Twan de Vries (and their co-authors) for making available a number of unpublished manuscripts. Finally, we thank Frans Wagenaar, Johan den Boon and Yvonne van der Meer for contributing to the illustrations for this review.

References


Bautista, E. M., Meuleenberg, J. J. M., Choi, C. S. & Molitor, T. W.


