

REVIEW

Nidovirus transcription: how to make sense...?

Alexander O. Pasternak, Willy J. M. Spaan and Eric J. Snijder

Molecular Virology Laboratory, Department of Medical Microbiology, Center of Infectious Diseases, Leiden University Medical Center, LUMC P4-26, PO Box 9600, 2300 RC Leiden, The Netherlands

Correspondence

Eric J. Snijder
e.j.snijder@lumc.nl

Many positive-stranded RNA viruses use subgenomic mRNAs to express part of their genetic information. To produce structural and accessory proteins, members of the order *Nidovirales* (corona-, toro-, arteri- and roniviruses) generate a 3' co-terminal nested set of at least three and often seven to nine mRNAs. Coronavirus and arterivirus subgenomic transcripts are not only 3' co-terminal but also contain a common 5' leader sequence, which is derived from the genomic 5' end. Their synthesis involves a process of discontinuous RNA synthesis that resembles similarity-assisted RNA recombination. Most models proposed over the past 25 years assume co-transcriptional fusion of subgenomic RNA leader and body sequences, but there has been controversy over the question of whether this occurs during plus- or minus-strand synthesis. In the latter model, which has gained considerable support recently, subgenomic mRNA synthesis takes place from a complementary set of subgenome-size minus-strand RNAs, produced by discontinuous minus-strand synthesis. Sense–antisense base-pairing interactions between short conserved sequences play a key regulatory role in this process. In view of the presumed common ancestry of nidoviruses, the recent finding that ronivirus and torovirus mRNAs do not contain a common 5' leader sequence is surprising. Apparently, major mechanistic differences must exist between nidoviruses, which raises questions about the functions of the common leader sequence and nidovirus transcriptase proteins and the evolution of nidovirus transcription. In this review, nidovirus transcription mechanisms are compared, the experimental systems used are critically assessed and, in particular, the impact of recently developed reverse genetic systems is discussed.

The increasing complexity of the nidovirus group

Nidoviruses are a group of enveloped positive-stranded RNA viruses. Currently known representatives mostly infect mammals (coronaviruses, toroviruses and arteriviruses), but do also have avian (coronaviruses) or invertebrate (roniviruses) hosts. Nidoviruses cause a variety of diseases, the outcome of which can range from an asymptomatic, persistent carrier-state to a sometimes fatal infection. The severity of coronavirus infection is exemplified by severe acute respiratory syndrome (SARS) in humans, which was caused by a newly emerged coronavirus that gripped worldwide attention in 2003 (Drosten *et al.*, 2003; Ksiazek *et al.*, 2003; Peiris *et al.*, 2003). In the wake of the SARS outbreak, several other novel coronaviruses, including two that infect humans (van der Hoek *et al.*, 2004; Fouchier *et al.*, 2004; Woo *et al.*, 2005), were identified and added to the growing list of nidoviruses that were first characterized during the past two decades.

During that same period of time, the systematic sequence analysis of virus genomes has changed the face of virus taxonomy. With the rise of virus genetics and molecular virology, it has become clear that comparative sequence analysis will provide the most solid basis for future virus classification systems. In addition, common strategies underlying the organization of viral genomes and common mechanisms for the regulation of viral genome expression have been recognized and have strengthened the case for using a genetic basis for virus taxonomy.

The current order *Nidovirales* is a perfect example of these developments. Since 1996, it officially unites the families *Coronaviridae* (genus *Coronavirus* and genus *Torovirus*) and *Arteriviridae*, which were initially considered to be completely unrelated. More recently, the new family *Roniviridae* was included, expanding the order into the domain of invertebrate hosts (prawns). The taxon derives its name from the common nidovirus strategy to express all genes located downstream of the replicase gene from a 3' co-terminal nested set of subgenomic (sg) mRNAs (*nidus* in Latin means nest; Fig. 1). However, the most compelling reason for nidovirus unification was found in the large replicase gene itself. In phylogenetic analyses of key replicase domains, including the RNA-dependent RNA polymerase (RdRp) and helicase, different nidovirus subgroups were found to cluster, suggesting that they share a common ancestor (Gorbalenya *et al.*, 1989; Snijder *et al.*, 1990a; den Boon *et al.*, 1991; Cowley *et al.*, 2000; for reviews see Cavanagh, 1997; de Vries *et al.*, 1997; Gonzalez *et al.*, 2003; Snijder *et al.*, 2005).

Despite the proposed common ancestry, the full-length replicase polyprotein (pp1ab; see below) is only between 3000 and 4000 aa residues long in arteriviruses, but is at least 6500 aa residues long in the case of coronaviruses, toroviruses and roniviruses. This size difference highlights one of the striking dissimilarities between viruses united in the order *Nidovirales*: their genome size ranges from 12 to 16 kb for arteriviruses to 26 to 31 kb for coronaviruses, toroviruses and roniviruses. In addition, major differences in virion morphology and the properties of the structural proteins underline the 'split personality' of nidoviruses: evolutionarily related replicase genes and an apparently conserved framework for genome expression have become linked to seemingly unrelated gene sets encoding structural proteins that specify virions of very different size and architecture.

Some 15 years ago, the detection of 3' co-terminal nested sets of sg mRNAs in cells infected with the arterivirus *Equine arteritis virus* (EAV; de Vries *et al.*, 1990) and the torovirus Berne virus [now known as *Equine torovirus* (EToV); Snijder *et al.*, 1990b] provided the first indication for a connection to coronaviruses, a family whose intriguing mechanism for sg mRNA transcription* had attracted attention for quite some time. Transcription essentially is the functional connection between the homologous and non-homologous parts of the nidovirus genome. By directing the synthesis of sg mRNAs, the conserved replicase gene controls the expression of the variable set of downstream genes. Studies into the mechanism of nidovirus transcription have been based on a combination of biochemical, genetic and molecular biological approaches and were stimulated by the development of systems allowing mutagenesis of regulatory sequences. These studies have clearly benefited from the characterization of virus groups distantly related to coronaviruses, but were also complicated by it. For example, the recent finding that – in contrast to coronavirus and arterivirus sg mRNAs – sg transcripts of toroviruses and roniviruses do not have a common 5' leader sequence illustrates that many details of sg mRNA production remain to be explored. The purpose of this review is to compare and contrast regulatory RNA elements and the molecular mechanisms of transcription in different nidoviruses, and to discuss these in the context of the current models for transcription.

Prototypic nidoviruses and their life cycle

For obvious reasons, the best-studied nidoviruses are those that either cause disease in humans, rodents or livestock and/or are most easily amplified in cell culture. An example is the arterivirus prototype EAV, which has the smallest (known) arterivirus/nidovirus genome (12.7 kb). This was a clear advantage for the development in 1996 of a reverse genetics system based upon a full-length cDNA clone from which infectious RNA can be produced (van Dinten *et al.*, 1997). In the family *Coronaviridae* (Lai & Cavanagh, 1997; Lai & Holmes, 2001), *Murine hepatitis virus* (MHV) has been extensively studied in terms of its molecular biology and, in particular, its transcription mechanism, a research area that will certainly benefit from the availability of MHV-reverse genetics systems recently developed in the laboratories of Ralph

*When its definition is followed to the letter, the term 'transcription' (i.e. 'the process by which genetic information encoded in one strand of DNA is copied into a complementary RNA strand') does not apply to the synthesis of sg mRNAs by nidoviruses and other RNA viruses.

Nevertheless, there is a clear functional parallel (production of RNA templates for protein synthesis) and the term 'transcription' has been used in studies on coronavirus sg mRNA synthesis from the very start. Consequently, for the purpose of this review and regardless of the lack of a DNA template, we will use the term 'transcription' for the synthesis of sg plus strands (sg mRNAs). Genome amplification, which results in the production of a full-length mRNA, will be referred to as 'replication'.

Baric (Yount *et al.*, 2002), and Stuart Siddell and Volker Thiel (Coley *et al.*, 2005). *Transmissible gastroenteritis virus* (TGEV) of swine, for which Luis Enjuanes and co-workers developed the first coronavirus full-length cDNA clone (Almazan *et al.*, 2000), has also been studied in some detail. EToV (Snijder & Horzinek, 1993) is the sole torovirus representative, thus far, that can be amplified and studied in cultured cells. In the new family *Roniviridae* (Roni stands for rod-shaped nidovirus), only the genome of *Gill-associated virus* (GAV) (Cowley *et al.*, 2000, 2002) has been fully sequenced and it is this virus for which an initial analysis of transcription was performed. Our knowledge of nidovirus transcription is largely based on studies with the limited number of model viruses mentioned above. Also, given the relatively large evolutionary distances between nidovirus subgroups, it is clear that data should be interpreted with caution when translating them to distant nidovirus clusters. This is true in particular for arteriviruses, in view of the size difference between their genome/replicase gene and those of other nidoviruses. For these reasons, the descriptions and conclusions that follow cannot be considered as definitive for all nidoviruses.

Following entry and uncoating, translation of the nidovirus genome (Fig. 1) is initiated at the replicase ORF1a start codon. The replicase gene is comprised of two large open reading frames (1a and 1b) that are connected by a -1 ribosomal frameshift site. Ribosomal frameshifting, promoted and coordinated by specific RNA signals, results in the C-terminal extension of a relatively small fraction of ORF1a-encoded polypeptides (pp1a) with the ORF1b-encoded polypeptide, which includes the most conserved enzymic functions including the RdRp. The pp1a and pp1ab replicase polyproteins are co- and post-translationally processed by two to four autoproteases that reside in pp1a. The number of mature replicase products (non-structural proteins or nsps) ranges from 12 or 13 in arteriviruses to 16 in most coronaviruses (see Ziebuhr *et al.*, 2000; Snijder *et al.*, 2003 and references therein). Guided by a number of ORF1a-encoded subunits that contain hydrophobic domains, most of the nidovirus replicase proteins associate with modified intracellular membranes to form a membrane-bound complex for RNA synthesis (for reviews see Lai & Holmes, 2001; Siddell *et al.*, 2005 and references therein), a common feature of animal positive-strand RNA viruses.

Recognition of RNA signals near the 3' end of the viral genome by the 'RdRp complex' precedes the initiation of synthesis of full-length minus-strand RNA (or antigenome), which in turn is the template for the synthesis of novel genome RNA. For the latter process, recognition signals present near the 3' end of the antigenome must be used (for reviews see Lai & Holmes, 2001; Siddell *et al.*, 2005 and references therein). Viral RNA synthesis is asymmetric and produces much more plus- than minus-strand RNA (Sawicki *et al.*, 2001). In addition to being utilized for boosting replicase expression and its own amplification, the newly synthesized nidovirus genome is also believed to be the template for the production of sg-length minus strands that are used as templates for transcription (see below). The incoming genomic RNA may also directly serve as a template for the production of sg-length minus strands, as it was possible to experimentally uncouple transcription from replication (Schelle *et al.*, 2005). Elegant biochemical studies by Stanley and Dorothea Sawicki (Sawicki *et al.*, 2001), using the coronavirus MHV as a model, showed that both antigenome and sg-length minus strands are

produced very early in infection. Each sg mRNA is produced from a corresponding transcription intermediate that contains the sg-length minus-strand template. These complexes synthesize the various sg mRNAs in non-equimolar, but relatively constant amounts. Also the ratio of the synthesis of genome to sg mRNAs is constant throughout the replication cycle (Sawicki *et al.*, 2001; Sawicki & Sawicki, 2005).

Nidovirus sg transcripts serve to express the structural proteins (and – in the case of coronaviruses – a variety of accessory proteins) from genes located in the 3'-proximal third of the genome, which are not accessible for ribosomes engaged in genome translation (Fig. 1). Although each mRNA, with the exception of the smallest one, is structurally polycistronic, they are functionally monocistronic and, with some exceptions, only the 5'-proximal ORF is translated. Ultimately, the newly synthesized genomes are encapsidated by the nucleocapsid (N) protein and progeny virions acquire their envelope by budding into the lumen of membranes of the endoplasmic reticulum-to-Golgi pathway (for a general overview of the nidovirus life cycle see Snijder *et al.*, 2005 and references therein).

The structure of nidovirus sg mRNAs: having a leader or not...

A principal feature of members of *Nidovirales*, although not unique for this virus order alone, is the generation of an extensive 3' co-terminal nested set of mRNAs from which the 3'-proximal region of the polycistronic genome is expressed. It was an early discovery that coronavirus sg transcripts are not only 3' co-terminal but also contain a common 5' leader sequence of about 65–100 nt, which is derived from the 5' end of the genome (Lai *et al.*, 1982a, 1983; Spaan *et al.*, 1982; Fig. 1). In some respects their mosaic nature resembles that of eukaryotic mRNAs generated by splicing. However, coronavirus replication occurs in the cytoplasm of the infected cell, and consequently the generation of sg mRNAs by fusion of sequences that are not contiguous in the genome was an unexpected feature. UV transcription inactivation studies essentially ruled out that the bulk of coronavirus sg RNA molecules were formed post-transcriptionally by *cis*-splicing of a genome-length precursor molecule (Jacobs *et al.*, 1981; Stern & Sefton, 1982; Yokomori *et al.*, 1992).

Following the identification of a short conserved sequence (later termed 'transcription-regulating sequence', TRS; Fig. 1), which is located at (i) the 3' end of the common leader, (ii) the 5' end of each mRNA 'body' segment, and (iii) the leader-to-body fusion site in the sg mRNA, it became apparent that base pairing between plus- and minus-sense copies of this regulatory sequence might direct the co-transcriptional fusion of sg mRNA leader and body in a process of discontinuous transcription (Baric *et al.*, 1983; Spaan *et al.*, 1983; Lai *et al.*, 1984). The initial studies on arterivirus transcription essentially corroborated these findings (de Vries *et al.*, 1990; den Boon *et al.*, 1995, 1996), although the arterivirus leader sequence and TRS were substantially longer and shorter, respectively, than the corresponding elements in coronaviruses.

Although a preliminary analysis of torovirus mRNAs (EToV) had already suggested the absence of common 5'-terminal sequences (Snijder *et al.*, 1990b), it was the analysis of

ronivirus sg transcripts and the more detailed analysis of EToV transcription, which firmly established that, apparently, discontinuous transcription is not a universal feature of nidoviruses. A leader sequence derived from the genomic 5' end was not found on the sg mRNAs of the ronivirus GAV (Cowley *et al.*, 2002). Even more surprising was the outcome of a thorough study on the 5'-terminal sequences of EToV mRNAs by Raoul de Groot and co-workers: whereas sg mRNAs 3, 4 and 5 of this virus lack a common 5' end, sg mRNA2 was shown to possess a short 5' leader sequence identical to the 5' end of the genome (van Vliet *et al.*, 2002).

The discontinuous step in coronavirus and arterivirus transcription: during plus- or minus-strand RNA synthesis?

Over the years, different models have been proposed to explain the fusion of the common 5' leader sequence to the different 3' body segments present in arterivirus and coronavirus sg mRNAs (previously reviewed by e.g. Sawicki & Sawicki, 1995, 2005; van der Most & Spaan, 1995; Lai & Cavanagh, 1997; Brian & Spaan, 1997; Snijder & Meulenberg, 1998; Lai & Holmes, 2001). Almost all models assume co-transcriptional fusion of leader and body, but the controversy whether the discontinuous step in RNA synthesis operates during plus- or minus-strand synthesis has kept the field busy and divided for years. In the two most prominent, but opposing models (commonly referred to as 'leader-primed transcription' and 'discontinuous extension of minus-strand RNA synthesis'), the TRS elements play a key role (Fig. 2). Both models include a base-pairing interaction between the sense copy of the TRS in the genomic leader (leader TRS) and the antisense copy of the TRSs present at the 5' end of each of the sg mRNA body segments (anti-body TRS). According to the 'leader-primed transcription' model (Fig. 2a; Baric *et al.*, 1983; Spaan *et al.*, 1983; Lai *et al.*, 1984), transcription is initiated from the 3' end of the antigenome to produce a leader primer of which the 3'-terminal leader TRS would base pair to the various anti-body TRSs in the antigenome. Subsequently, the leader transcript would be extended to complete the sg mRNA. Thus, this model proposes that the discontinuous step takes place during plus-strand synthesis and that the body TRS complements in the antigenome essentially act as promoters for transcription.

The leader-primed transcription model was largely based on the inability to detect sg-length minus strands in coronavirus-infected cells (Lai *et al.*, 1982b). However, the subsequent discovery of such molecules, for both coronaviruses and arteriviruses (Sethna *et al.*, 1989; Sawicki & Sawicki, 1990; den Boon *et al.*, 1996), sparked the conception of alternative models. David Brian and colleagues, upon first identifying a nested set of sg minus-strand RNAs in TGEV-infected cells (Sethna *et al.*, 1989), proposed the sg-length minus strands to be intermediates in the amplification of sg mRNAs. Thus each sg mRNA, initially produced by leader-primed transcription or possibly even present in the incoming virus particle, would establish its own replicon to boost sg mRNA accumulation. However, for both coronaviruses (Masters *et al.*, 1994) and arteriviruses (Tijms *et al.*, 2001), it was later shown that genomic sequences downstream of the leader TRS (and therefore lacking in all sg mRNA species) are required for RNA replication.

Subsequently, Stanley and Dorothea Sawicki (Sawicki & Sawicki, 1995) proposed discontinuous extension of minus-strand RNA synthesis as an alternative (Fig. 2b). Not plus-, but minus-strand sg RNA synthesis was proposed to be discontinuous, with attenuation of RNA synthesis occurring in the different body TRS regions of the genomic template. The nascent sg-length minus strand, having an anti-body TRS at its 3' end, would then base pair with the leader TRS, be completed by extension with the anti-leader sequence and subsequently serve as template for transcription. In recent years, this model has gained considerable experimental support from both biochemical and genetic studies (van Marle *et al.*, 1999a; Baric & Yount, 2000; de Vries *et al.*, 2001; Sawicki *et al.*, 2001; Pasternak *et al.*, 2001; Zuniga *et al.*, 2004).

It should be noted that leader-priming and attenuation of minus-strand synthesis at body TRSs are not mutually exclusive. Formally, the possibility that sg mRNAs are formed by leader-primed transcription from the attenuated sg-length minus-strand templates cannot be ruled out (van der Most *et al.*, 1994).

Attenuation of minus-strand RNA synthesis: the common step in nidovirus transcription?

The transcription mechanism involving a discontinuous step, as described above for coronaviruses and arteriviruses, is unique and contrasts with the mechanism used by many other positive-strand RNA viruses that produce sg mRNAs by internal initiation from 'promoters' in the antigenome (Miller *et al.*, 1985; reviewed by Miller & Koev, 2000). However, similarities may exist to the mechanism used by a smaller group of viruses, exemplified by *Flock house virus*, *Tomato bushy stunt virus* and *Red clover necrotic mosaic virus*, that employ premature termination (attenuation) of minus-strand RNA synthesis (reviewed by White, 2002). These viruses also produce sg-length minus strands that serve as templates for transcription, but discontinuous extension of the nascent minus-strand RNA does not occur.

Strikingly, this is exactly the mechanism that may apply to the transcription of ronivirus sg mRNAs and all but the largest of the torovirus sg mRNAs (Fig. 3). Although reports on the minus-strand RNAs (genome length or sg-length) produced by these two virus groups have not yet been published, one might speculate that minus-strand RNA synthesis is attenuated, as in coronaviruses and arteriviruses, to produce sg-length templates for transcription. This would explain the conservation in torovirus and ronivirus genomes of the apparent equivalents of the body TRSs found in coronaviruses and arteriviruses (Snijder *et al.*, 1990b; Cowley *et al.*, 2002; van Vliet *et al.*, 2002). Instead of extension of the nascent minus-strand with the anti-leader sequence (Fig. 3a), roniviruses and toroviruses may directly use the attenuated minus-strand products as templates for transcription (Fig. 3b). Consequently, if sg-length minus strands indeed exist in torovirus- and ronivirus-infected cells, it may be the attenuation step during minus-strand RNA synthesis that is the common denominator in nidovirus transcription.

RNA replicons and full-length cDNA clones as tools in nidovirus transcription research

So far, most studies on nidovirus transcription have been done using coronaviruses, the virus group for which discontinuous transcription was first proposed. However, until quite recently, reverse genetics systems based on full-length cDNA clones were not available for coronaviruses (see below). Consequently, studies were based on the use of defective interfering (DI) RNA-derived replicons that contained one or more natural or synthetic (inserted) body TRSs. Such replicon RNAs, generated *in vitro* from cloned cDNA templates, can replicate and engage in transcription when transfected into cells that are infected with the parental virus ('helper virus') to provide the replicase proteins required for RNA synthesis.

Although representative to a certain extent, such replicon systems may only 'reflect' the transcriptional processes as they occur in the coronavirus-infected cell. The first and most important drawback is the generally low transcriptional activity of DI RNA-based body TRSs. The body TRSs in the coronavirus DI RNA-based replicon systems reported to date were generally unable to promote the transcription of amounts of sg mRNA comparable with those produced from the same body TRSs residing in the full-length genome of the parental virus. In fact, most DI RNA-based body TRSs produce one or two orders of magnitude less sg mRNAs. Consequently, it is not entirely clear if and how the conclusions obtained with such replicon systems can be extrapolated to transcription from the full-length virus genome. Whereas most point mutants of an MHV body TRS introduced into a DI RNA-based replicon supported transcription levels approaching those of the original TRS (Joo & Makino, 1992; van der Most *et al.*, 1994), TRS point mutations could decrease transcription about 100-fold in the context of the EAV full-length cDNA clone system (van Marle *et al.*, 1999a; Pasternak *et al.*, 2001) and up to 1000-fold in the TGEV infectious clone system (see below; Zuniga *et al.*, 2004). Secondly, due to the high frequency of RNA recombination with the helper virus genome in coronavirus DI RNA replicon systems, it was not possible to test the influence of leader TRS mutations on transcription, a problem that was readily overcome by the development of full-length molecular clones (van Marle *et al.*, 1999a; Pasternak *et al.*, 2001, 2003; Zuniga *et al.*, 2004). Third, although a number of studies have addressed the effects of flanking sequences on body TRS activity in DI RNA replicon systems, the overall structural context of these regulatory sequences, and not just the immediate flanking sequences, may be important for their transcriptional activity (see below).

An alternative to the DI RNA replicon-based studies of coronavirus transcription was developed by Paul Masters and co-workers. Targeted homologous recombination was used to introduce an additional body TRS into the MHV genome (Hsue & Masters, 1999). However, compared with modification of an infectious cDNA clone, this system is quite laborious, also it does not allow modification of the genomic leader and its TRS, and has the important disadvantage that only viable recombinants can be analysed since recombinant virus needs to be grown and passaged prior to analysis. Consequently, a first cycle analysis of transcription is not possible, making this system impractical to e.g. support large-scale TRS mutagenesis studies.

Clearly, the construction of full-length cDNA clones for different nidoviruses (van Dinten *et al.*, 1997; Meulenber *et al.*, 1998; de Vries *et al.*, 2001; Almazan *et al.*, 2000; Thiel *et al.*, 2001; Casais *et al.*, 2001; Yount *et al.*, 2002, 2003; Coley *et al.*, 2005) was a major step towards development of more straightforward systems for transcription research. The reverse genetics system for the arterivirus EAV, for example, permits direct transfection of *in vitro* synthesized EAV full-length RNA into BHK-21 cells with reasonable efficiency, permitting first cycle analysis of viral RNA synthesis and rapid screening for specific virus phenotypes using conventional biochemical assays. Full-length cDNA copies of the TGEV genome were used for the first extensive mutagenesis studies targeting coronavirus transcription in the context of the full-length genome (Zuniga *et al.*, 2004; Curtis *et al.*, 2004; Sola *et al.*, 2005). Although the currently available systems for coronavirus reverse genetics frequently require additional passaging of mutant viruses before analysis, this problem has been circumvented by targeting the TRSs of sg mRNAs whose translation product is not essential for virus replication in cell culture.

The TRS: key regulatory element in nidovirus transcription

The nidovirus TRS is an AU-rich element that was often called intergenic sequence or leader-body junction site in the older literature on coronaviruses and arteriviruses, respectively. The extent of sequence identity between the coronavirus leader TRS region and the different body TRSs ranges from 7 to 18 nt, whereas the corresponding sequence of arteriviruses usually is between 5 and 8 nt long. In toroviruses, a conserved 12 nt long sequence element is present upstream of ORFs 3, 4 and 5. Although, as explained above, there are important mechanistic differences with coronaviruses and arteriviruses, it was recently shown that a 16 nt cassette containing this torovirus TRS can direct the transcription of an sg mRNA when inserted into an EToV DI RNA-based replicon (Smits *et al.*, 2005).

As discussed above, in the 1990s, similar DI RNA-based systems formed the platform for the initial mutagenesis studies aimed at understanding the function of coronavirus body TRSs and their regulation. Shinji Makino and co-workers (Makino *et al.*, 1991) inserted the MHV RNA7 body TRS and its flanking sequences into a DI RNA replicon resulting in the transcription of a DI RNA-derived sg mRNA. This system was used for site-directed mutagenesis of body TRS and flanking sequences to delimitate *cis*-acting elements regulating transcription. Using this system, Joo & Makino (1992) identified the leader-body fusion sites of DI RNA-derived sg mRNAs. The MHV body TRSs are 5'-AAUCUAAAC-3' (or a closely related sequence), whereas the 3' end of the leader contains two to four 5'-UCUAA-3' repeats (depending on the virus strain; Makino *et al.*, 1988). The mRNAs were found to contain two to four 5'-UCUAA-3' repeats, depending on alternative base-pairing possibilities within the duplex (Makino *et al.*, 1988). Joo & Makino (1992) found that the first 5'-UCUAA-3' sequence was contributed by the leader and that leader-to-body fusion most likely took place at the first or the second nucleotide of the second repeat. Using a similar system, van der Most *et al.* (1994) concluded that the leader-body junction occurs at multiple sites within the duplex with a preference for 3'-proximal nucleotides within the

body TRS. The same authors showed that the contribution of the leader TRS to the sg mRNA can be as little as 3 nt. At that time, these data were interpreted in the context of the leader-primed transcription model and a variant involving a 'back-trimming' mechanism (Baker & Lai, 1990), which was proposed to employ a 3'→5' nuclease to trim the leader sequence prior to its extension with the mRNA body.

Studies using the EAV full-length cDNA clone system (van Marle *et al.*, 1999a; Pasternak *et al.*, 2001, 2003) have rigorously demonstrated that transcription depends on leader TRS-body TRS duplex formation (Fig. 4), and have yielded starting information concerning additional RNA and protein determinants of arterivirus transcription. More recently these conclusions were corroborated for the coronavirus TGEV using mutagenesis of a full-length cDNA clone (Zuniga *et al.*, 2004). It was shown that disruption of the base-pairing interaction dramatically affects transcription and that the introduction of compensatory mutations could restore activity (see also below and Fig. 4). Generally, these studies revealed that the relative amount of sg mRNA in both arteriviruses and coronaviruses correlates with the calculated stability of the corresponding leader TRS-body TRS duplex. However, this is clearly not the only factor determining sg mRNA amounts. Sequences flanking the core TRS were also shown to influence transcription in the context of the TGEV full-length genome (Curtis *et al.*, 2004; Sola *et al.*, 2005), confirming the earlier results obtained with DI RNA replicons (van der Most *et al.*, 1994; Jeong *et al.*, 1996; An & Makino, 1998; Ozdarendeli *et al.*, 2001; Alonso *et al.*, 2002). In addition to the effect of flanking sequences, it should be noted that when core TRS and TRS-flanking sequences were standardized by head-to-tail insertion of several copies of a body TRS-containing cassette into an EAV full-length clone-derived replicon, a perfect gradient of sg mRNA abundance, progressively favouring smaller RNA species, was observed (Pasternak *et al.*, 2004). This confirmed earlier theoretical work on the MHV genome (Konings *et al.*, 1988) and studies with coronavirus DI RNA replicon systems (van Marle *et al.*, 1995; Joo & Makino, 1995; Krishnan *et al.*, 1996), which have shown that relative order and/or location of TRSs in the genome play an important role.

An additional aspect that was investigated was the origin of the TRS that is present at the leader-body junction in the sg mRNA. Whereas the leader-primed transcription model (Fig. 2a) predicts this sequence to be a copy of the leader TRS, the discontinuous minus-strand extension model (Fig. 2b) predicts exactly the opposite, i.e. the TRS copy in the mRNA should be derived from the body TRS. Studies with MHV DI RNA replicons demonstrated that the leader-body junction could be partially derived from the leader TRS and partially from the body TRS (van der Most *et al.*, 1994). Using EAV TRS mutants with residual transcriptional activity, it was demonstrated that the body TRS could specify the entire sequence of the TRS copy in the sg mRNA, thus providing strong support for the discontinuous minus-strand extension model (Pasternak *et al.*, 2001). This was subsequently confirmed also for coronaviruses (Zuniga *et al.*, 2004; Hussain *et al.*, 2005). Additional evidence in favour of the discontinuous minus-strand extension model stems from the fact that the aberrant transcripts originating at non-canonical body TRS-like sequences of non-viral origin had their entire leader-body junction sequences derived from the 'body TRS' (de Vries *et al.*, 2001; Sola *et al.*, 2003).

Cis-acting signals directing nidovirus transcription: primary sequence or higher order RNA structure?

At first glance, the sequence conservation of nidovirus body TRSs might suggest sequence-specific recognition by a protein factor. Indeed, the RdRps of some positive-strand RNA viruses are able to bind to sg promoters in a sequence-specific manner, the best studied example being the RdRp of *Brome mosaic virus* (BMV) (Siegel *et al.*, 1997, 1998; Adkins *et al.*, 1997; Stawicki & Kao, 1999). However, BMV transcription is initiated internally on the genome-length negative strand (Miller *et al.*, 1985), a mechanism obviously requiring distinct *cis*-acting signals and *trans*-acting factors, but clearly different from the transcription models proposed for nidoviruses.

In recent years, an increasing number of base-pairing interactions, often spanning more than 1000 nt, have been implicated in transcription or replication of various bacterial, plant and animal viral systems (Klovins *et al.*, 1998; Zhang *et al.*, 1999; Kim & Hemenway, 1999; Choi *et al.*, 2001; Lindenbach *et al.*, 2002; Choi & White, 2002; Lin & White, 2004). In one 'extreme' case, transcription was found to require an intermolecular RNA–RNA interaction (Sit *et al.*, 1998). Interestingly, compensatory mutations simultaneously introduced in the distal and proximal regulatory elements of the *Potato virus X* could restore transcription, but not to wild-type levels, indicating that both this long distance RNA–RNA interaction and the sequence of the proximal elements are involved (Kim & Hemenway, 1999).

The latter situation is strikingly similar to the results obtained during leader TRS-body TRS co-variation mutagenesis studies with EAV (Pasternak *et al.*, 2001) and TGEV (Zuniga *et al.*, 2004). EAV studies provided genetic evidence that body TRSs have a specific function in transcription distinct from the formation of a duplex with the leader TRS (Pasternak *et al.*, 2001). Transcription defects caused by a subset of body TRS point mutations could not be restored by introduction of the compensatory mutation in the leader TRS (Fig. 4b). Strikingly, despite their primary structure conservation, these body TRS-specific requirements apparently differ between body TRSs. For example, the EAV RNA6 body TRS contains a C at position 6, whereas this nucleotide is absolutely not tolerated at the same position in the RNA7 body TRS. Moreover, whereas the U₁A substitution in the latter allowed for 40 % of mRNA7 transcription and could not be rescued by the compensatory leader TRS mutation (Pasternak *et al.*, 2001), the same substitution in the RNA6 body TRS caused an almost complete shutdown of transcription, which could be efficiently rescued in the double mutant (Fig. 4c; D. D. Nedialkova, A. O. Pasternak & E. J. Snijder; unpublished data). Also the fact that a mutant carrying five mutations (5'-UCAACU-3'→5'-AGUUGU-3') in leader TRS and RNA7 body TRS could support a certain level of mRNA7 transcription argues against sequence-specific recognition of the body TRS by a protein factor (van Marle *et al.*, 1999a). However, this does not exclude the involvement of a regulatory protein factor (Pasternak *et al.*, 2004) that would recognize e.g. higher order, rather than primary, structural motifs in body TRS regions. Furthermore, whereas it was found that the extent of base pairing and the stability of the leader TRS-body TRS duplex play an important

role in regulation of transcription (Pasternak *et al.*, 2003; Zuniga *et al.*, 2004; Sola *et al.*, 2005), some of the leader TRS mutations tested for EAV (Pasternak *et al.*, 2001) were much more deleterious than would be expected under the assumption that they only destabilized the duplex (see Pasternak *et al.*, 2003, for details). Although, according to the 'free-energy threshold concept' proposed for TGEV (Zuniga *et al.*, 2004; Sola *et al.*, 2005), a minimum free-energy value of leader TRS-body TRS duplex may be required to promote strand transfer, the transcription mechanism may require base pairing of central ('core') nucleotides of the TRS to properly position the nascent strand on the template before elongation. Hence, conservation of the TRS primary structure may serve to guarantee the fidelity of the strand transfer process, which could be driven by protein factors bound to secondary structure motifs in the leader TRS and/or body TRS regions.

Secondary structure predictions of the arterivirus leader region (van Marle *et al.*, 1999a; van den Born *et al.*, 2004) place the leader TRS in a hairpin loop structure referred to as the 'leader TRS hairpin' or LTH (Fig. 5a). A similar conformation can be predicted for most coronaviruses and previously the existence of a leader TRS presenting hairpin in *Bovine coronavirus* was firmly supported by structure probing experiments (Chang *et al.*, 1996). Recent studies, in which the EAV LTH was duplicated to allow its mutagenesis without affecting replication and genome translation, support a critical role for this domain in transcription. The leader TRS duplicate yielded novel sg mRNAs with significantly extended leaders. Furthermore, a construct with two functional LTHs was able to produce a perfect double-nested set of sg mRNAs (van den Born *et al.*, 2005).

For EAV body TRS regions, it was hypothesized (van Marle *et al.*, 1999a; Pasternak *et al.*, 2000, 2001) that *cis*-acting signals would at least partially have the form of higher order RNA structures, either in the plus-strand template or in the nascent minus-strand. An important argument in favour of such regulatory structural motifs is the existence of 'potential body TRSs' in arterivirus genomes that match the leader TRS precisely but still either do not function as body TRS at all or produce only minute amounts of sg mRNAs (den Boon *et al.*, 1996; van Marle *et al.*, 1999a; Pasternak *et al.*, 2000). A striking example is TRS4.1 of EAV, which produces 40 times less sg mRNA than the adjacent TRS4.2 (Pasternak *et al.*, 2000), despite its 8 nt match with the leader TRS, compared with a 6 nt match for TRS4.2. It could be argued that protein complexes bound to the downstream TRS4.2 prevent the formation of the same complex on TRS4.1, which is located only 17 nt upstream. Indeed, downstream body TRSs were shown to suppress the activity of upstream TRSs when placed close to each other (less than 100 nt) in the coronavirus replicon system (Joo & Makino, 1995; Krishnan *et al.*, 1996). However, it is questionable whether this notion can be applied to EAV, since TRSs 3.1 and 3.2 produce equal amounts of sg mRNA (Pasternak *et al.*, 2000) despite also being quite close to one another (32 nt distance).

Transcription of the largest sg mRNA of toroviruses, which carries a short leader derived from the genomic 5' end, was proposed to involve an RNA structure in the body TRS region. An RNA hairpin is predicted to be present in the genomic template, just upstream of the proposed attenuation site for minus-strand synthesis (Fig. 5b). A short piece of sequence at this position is

identical to a sequence in the genomic 5'-proximal region and appears to serve as a crossover site for the acquisition of the 18 nt leader sequence (van Vliet *et al.*, 2002). RNA structure predictions have also been made for the body TRS regions of EAV (Pasternak *et al.*, 2000) and of *Bovine coronavirus* (Ozdarendeli *et al.*, 2001). Remarkably, both studies predicted the 'active' TRSs in the plus-strand genomic template to be located in non-base-paired regions, whereas the less active or non-active TRSs are predicted to be completely or partially base paired. None of these predictions, however, could be verified experimentally (Ozdarendeli *et al.*, 2001; Pasternak, 2003).

In addition to leader and body TRSs, the 3' end of the viral genome is obviously crucial for both transcription and replication. Although based on data obtained with DI RNA replicons, it was concluded that ~300 nt at the MHV genomic 3' end are required for transcription (Lin *et al.*, 1996), whereas only the 3'-proximal 55 nt of the same domain [and the poly(A) sequence] are required for initiation of negative-stranded RNA synthesis (Lin *et al.*, 1994). This difference suggests the presence of a transcription-specific *cis*-acting sequence at the 3' end of the viral RNA.

Viral proteins involved in nidovirus transcription

Components of the nidovirus replicase are obvious candidates for a role as regulatory factors in transcription. In addition to the core proteins of the RNA-synthesizing machinery, accessory protein functions specifically involved in transcription may have evolved that are either common to nidoviruses or specific for certain nidovirus subgroups. In the case of arteriviruses, nsp1 and nsp10 have been implicated in transcription (van Dinten *et al.*, 1997; van Marle *et al.*, 1999b; Tijms *et al.*, 2001; Tijms, 2004; Seybert *et al.*, 2005). Remarkably, both proteins contain (putative) zinc-binding domains, which may facilitate RNA-protein interactions (van Dinten *et al.*, 2000; Tijms *et al.*, 2001). Mutations in the conserved putative zinc-binding domain of nsp1 either abolish transcription completely, without dramatic effects on genome replication or seem to influence the balance between replication and transcription (Tijms *et al.*, 2001; M. A. Tijms, J. C. Zevenhoven, A. E. Gorbalenya & E. J. Snijder, unpublished data). Interestingly, zinc finger structures in the HIV N protein facilitate nascent DNA-strand transfer between RNA templates (Guo *et al.*, 2000). A zinc finger domain in the p23 protein of *Citrus tristeza virus* (CTV), a plant closterovirus with a genome organization remarkably similar to that of nidoviruses, was shown to mediate the activity of this protein in asymmetrical RNA accumulation in CTV-infected plants (Satyanarayana *et al.*, 2002).

Structure-function studies of the coronavirus replicase were promoted by the recent SARS outbreak. Bioinformatic analyses by Alexander Gorbalenya and others identified five domains with distant relationships to cellular enzymes involved in RNA metabolism (Snijder *et al.*, 2003; von Grotthuss *et al.*, 2003; Yan *et al.*, 2003). Particularly relevant to viral RNA synthesis was the analysis of an array of conserved domains in the C-terminal region of pp1ab (now named nsp14, nsp15 and nsp16), which were predicted to possess 3'→5' exonuclease (ExoN), uridylate-specific endoribonuclease (NendoU) and S-adenosylmethionine-dependent

ribose 2'-O-methyltransferase (2'-O-MT) activities. Strikingly, all three domains are conserved in coronaviruses, toroviruses and roniviruses, but only the NendoU domain is also found in arteriviruses (nsp11). It is this domain, which may function as a homohexamer (Guarino *et al.*, 2005), for which experimental studies recently confirmed its Mn²⁺-dependent *in vitro* endoribonuclease activity and importance for virus replication (Ivanov *et al.*, 2004; Bhardwaj *et al.*, 2004). Although the precise determinants of NendoU specificity remain to be determined, a preference for cleavage at specific uridylate-containing sequences in dsRNA was reported and an interesting hypothesis of the involvement of NendoU in transcription was formulated (Ivanov *et al.*, 2004). In the EAV-reverse genetics system, NendoU mutagenesis exerted pleiotropic effects on viral RNA synthesis (Posthuma *et al.*, 2006). While some mutations rendered RNA synthesis undetectable, others induced only a moderate reduction, with sg RNA synthesis consistently being more strongly affected than genome replication.

In coronaviruses, one or more of the small replicase subunits from the C-terminal region of pp1a (nsp7–10), which do not appear to have a counterpart in arteriviruses, may be directly or indirectly involved in viral RNA synthesis. The MHV RdRp-containing subunit (nsp12) was proposed to interact with nsp8 and nsp9 (Brockway *et al.*, 2003). Structural and biochemical studies characterized SARS-CoV nsp9 as a single-stranded RNA-binding protein displaying a new variant of the OB-fold (Egloff *et al.*, 2004; Sutton *et al.*, 2004). Furthermore, SARS-CoV nsp7 and nsp8 were recently reported to form a unique hexadecameric structure that is probably capable of encircling RNA and may operate as an accessory factor for the RdRp complex (Zhai *et al.*, 2005). Recently, nsp10 was also implicated in viral RNA synthesis, on the basis of the phenotype of an MHV mutant carrying a temperature-sensitive mutation in this subunit (Sawicki *et al.*, 2005; see below).

The latter elegant study by Stanley Sawicki, Stuart Siddell and colleagues also underlines that, in addition to reverse genetics, classical 'forward' genetics, based on the use of virus mutants with a conditional defect in RNA synthesis, is a powerful approach to dissect the complex web of protein interactions that governs nidovirus RNA synthesis. Using MHV, the only nidovirus for which a substantial set of temperature-sensitive mutants is available (Sturman *et al.*, 1987; Schaad *et al.*, 1990; Sawicki *et al.*, 2005), mutations in six different nsps were identified (nsp4, nsp5, nsp10, nsp12, nsp14 and nsp16), which – under restrictive conditions – all interfered with the assembly of a functional RdRp complex. Different types of RNA synthesis defects were observed, including the inability to synthesize minus-strand RNA (nsp10 mutation) and the apparent inability to switch from minus- to plus-strand RNA synthesis (interestingly caused by a mutation in the C-terminal domain of nsp5, the viral main proteinase subunit). The detailed characterization of these mutant phenotypes, and the possibility to hunt for second site revertants, may provide valuable information on the role of different nidovirus nsps in RNA synthesis and key functional interactions within the RdRp complex.

Finally, the emerging role of the N protein in coronavirus RNA synthesis should be mentioned. In arteriviruses this role seems to be non-existent, since both replication and transcription are not affected by inactivation of N protein expression (Molenkamp *et al.*, 2000b; Pasternak *et al.*, 2001). The coronavirus N protein, on the other hand, has been implicated in

regulation of RNA synthesis for a long time and was, for example, reported to specifically interact with leader TRS in the case of MHV (Nelson *et al.*, 2000). Recent studies in reverse genetic systems revealed that the N protein is required for the efficient initiation of replication following transfection of *in vitro* generated *Human coronavirus 229E* (HCoV-229E) infectious RNA (Schelle *et al.*, 2005) or of a plasmid expressing a TGEV replicon (Almazan *et al.*, 2004). Strikingly, using a replicon system for HCoV-229E, it was found that replication, but not transcription, was impaired in the absence of the N protein, suggesting that this structural protein may be involved in regulating the balance between these two processes (Schelle *et al.*, 2005).

In general, whereas it can be hard to identify *cis*-acting higher order RNA structure motifs and to prove their function by reverse genetics (Pasternak, 2003), it should be possible to identify viral and/or host proteins that interact with body TRS regions. For example, yeast three-hybrid systems (Sengupta *et al.*, 1999) and biochemical assays in a cell-free environment can help to identify such factors, and their role may subsequently be confirmed *in vivo* by reverse genetics.

Coronavirus and arterivirus transcription: a variant of similarity-assisted RNA recombination?

The transcription mechanism described above for arteriviruses and coronaviruses, where both leader-body TRS duplex formation and a distinct function of the body TRS are required for proper transfer of the nascent strand, would resemble the mechanism of the typical similarity-assisted ('copy-choice') RNA recombination described for *Turnip crinkle virus* (TCV). In this system, both sequence similarity between the parental strands and an RdRp-binding hairpin in the acceptor strand are necessary for strand transfer (Nagy *et al.*, 1998; Nagy & Simon, 1998a, b; for a review see Nagy & Simon, 1997). Indeed, coronavirus and arterivirus discontinuous RNA synthesis has been proposed to resemble high-frequency copy-choice RNA recombination (see e.g. Spaan *et al.*, 1983; Chang *et al.*, 1996; Brian & Spaan, 1997; van Marle *et al.*, 1999a and references therein). In cells simultaneously infected with two different MHV strains, up to half of the mRNAs may carry a leader sequence from the co-infecting strain (Makino *et al.*, 1986), suggesting free-leader exchange during transcription and a mechanism that can join leader and body sequences derived from different templates (Zhang & Lai, 1994; Jeong & Makino, 1994). Later, Britton and colleagues demonstrated highly efficient leader switching during the rescue of defective RNAs by heterologous strains of the coronavirus *Infectious bronchitis virus*, suggesting that the discontinuous step may be a part of the normal DI RNA replication cycle (Stirrup *et al.*, 2000). Like retroviral reverse transcriptases (Peliska & Benkovic, 1992, 1994; Wu *et al.*, 1995), coronavirus and arterivirus RdRps may be especially prone to switch templates, producing recombinant genomes in the case of homologous recombination (Lai *et al.*, 1985; Keck *et al.*, 1988a, b; Liao & Lai, 1992; Molenkamp *et al.*, 2000a; Pasternak *et al.*, 2004) or sg RNAs in the special case of discontinuous transcription.

The production of minor sg mRNAs from non-canonical body TRS-like sequences of non-viral origin (Fischer *et al.*, 1997; de Vries *et al.*, 2001; Curtis *et al.*, 2002), as well as 'induction' of minor sg mRNAs produced from non-canonical TRSs upon mutagenesis of the leader TRS (Zhang & Lai, 1994; Pasternak *et al.*, 2003; Zuniga *et al.*, 2004) are consistent with a high template switching frequency of coronavirus and arterivirus polymerases. It has been reported that recombination in both virus groups occurs more frequently at the 3' end of the genome (Fu & Baric, 1992, 1994; Molenkamp *et al.*, 2000a), suggesting that body TRSs may be recombination hot spots. Likewise, *Luteovirus* sg promoters were proposed to be recombination hot spots (Koev *et al.*, 1999), and hairpins required for TCV recombination served as replication enhancers (Nagy *et al.*, 1999). This suggests that RNA recombination may be driven by similar factors and/or signals as discontinuous transcription in nidoviruses.

In several viral systems, 5' termini in donor templates and stable hairpin structures in both donor and acceptor templates have been shown to promote RNA recombination (Cascone *et al.*, 1990, 1993; Nagy & Bujarski, 1993; Carpenter *et al.*, 1995; White & Morris, 1995; Nagy *et al.*, 1998). Also, sg promoters in several positive-strand RNA viruses include RNA secondary structure motifs (Wang *et al.*, 1999; Koev *et al.*, 1999; Haasnoot *et al.*, 2000). It is not clear whether all such structures interact with specific regulatory proteins or whether they might just mechanistically promote template-switching or pausing. Nagy *et al.* (1995) have shown that mutations in the helicase-like domain of BMV protein 1a could alter the sites of RNA recombination. It is interesting that EAV nsp10, which also plays a specific role in transcription (see above), possesses helicase activity (Seybert *et al.*, 2000), although it remains to be studied whether this activity is directly involved in transcription.

(Dis)continuous transcription: to jump or not to jump?

The question whether arterivirus and coronavirus transcription is truly discontinuous (involving a 'jump' of nascent strand and/or RdRp from one template site to the other) or quasi-continuous, as suggested by Zhang *et al.* (1994), is an interesting unresolved issue. The latter mechanism was proposed to involve looping-out of the template and formation of a triple-stranded intermediate, possibly mediated by viral or host proteins, in a manner similar to DNA-dependent RNA transcription for which promoter and enhancer elements are also brought together by protein-protein interactions (Lai, 1998). In the first model, the RdRp complex is assumed to temporarily dissociate from template and possibly also nascent strand, whereas in the second model it remains associated with the RNA strands. In the first model, the body TRSs in the plus-strand template would behave as terminators of RNA synthesis, whereas in the second model they would only promote pausing. A key parameter in this debate is the processivity of the nidovirus RdRp complex, which is unknown. As in the case of copy-choice RNA recombination (Jarvis & Kirkegaard, 1991; Nagy & Simon, 1997), a non-processive polymerase would tend to dissociate from the template, whereas a processive one would be more likely to remain associated with it. The non-processive model for RNA recombination, as proposed for TCV (Cascone *et al.*, 1990, 1993), implies that donor or nascent strand would

contain termination signals and the acceptor strand would contain signals for retargeting of RdRp and nascent strand. Indeed, the acceptor strand in the TCV satellite RNA recombination system contains a hairpin structure that was shown to bind the RdRp and recruit the nascent strand (Nagy *et al.*, 1998; Nagy & Simon, 1998a, b). Leader TRS-containing hairpin structures, as proposed for both arterivirus and coronaviruses (Chang *et al.*, 1996; van Marle *et al.*, 1999a; van den Born *et al.*, 2004), may fulfil a similar role during transcription of sg mRNAs. The termination of nascent strand synthesis, in turn, may be mediated by primary or higher order RNA structure of donor or nascent strand, in a manner similar to the *rho*-independent transcription termination in *Escherichia coli* (Wilson & von Hippel, 1995) or the eukaryotic RNA polymerase I termination system (Reeder & Lang, 1997). Alternatively, as proposed for RNA recombination in BMV or poliovirus (Jarvis & Kirkegaard, 1991; Nagy & Bujarski, 1995, 1996), the strand transfer process may be processive, with the polymerase occasionally pausing and sliding backwards, which would result in the extrusion of the 3' end of the nascent strand from the RdRp complex. This 3' end would then be free to hybridize to a complementary region on the acceptor template and subsequently RNA synthesis could be resumed. This mechanism would resemble the transcriptional arrest in *E. coli* (Nudler *et al.*, 1997; Komissarova & Kashlev, 1997a, b). Interestingly, Nudler *et al.* (1996) have shown that *E. coli* RNA polymerase can switch DNA templates in a processive manner and that this processivity depends on the zinc finger structure in the N terminus of the β' polymerase subunit. This suggests that the basic features of RNA synthesis, including the ability of the enzyme to continuously slide forward and backward on the nascent strand and to switch templates, may be conserved at least between RNA viruses and bacteria. The ability of the BMV RdRp to initiate transcription from DNA templates (Siegel *et al.*, 1999; Sivakumaran & Kao, 1999) supports this concept of evolutionary conservation. Thus, it is possible, for example, that specific factors that mediate strand transfer in extensively studied DNA recombination systems (e.g. RecA) may either have analogues among viral proteins or even function directly as host factors required for viral RNA recombination and/or transcription. In the former case, the basic recombination machinery, just as the machinery involved in RNA synthesis, would be conserved between viruses and higher organisms. In the latter case, viruses could have partially adapted to features of the host recombination machinery.

Exploring the mechanistic details of nidovirus transcription

Considerable genetic and biochemical support has been obtained for the discontinuous minus-strand extension model (Sawicki & Sawicki, 1995) in which body TRSs are viewed as attenuators of minus-strand synthesis. However, strictly speaking, this model will only be proven when the production of specific reaction intermediates [e.g. nascent (attenuated) minus strands not yet containing the anti-leader sequence] can be analysed. Thus far, conditions/mutations that allow the accumulation of transcription intermediates or the uncoupling of sg minus- and plus-strand synthesis in infected cells have not been identified. The theoretical possibility that

sg-length minus strands could be copied from sg mRNAs and vice versa further complicates experimental studies.

Site-directed mutagenesis of RNA or protein sequences, in combination with sophisticated biochemical assays, may allow the selective and sensitive detection of intermediates that could arise when a specific stage of discontinuous transcription would be blocked. For example, a mutation in the EAV leader TRS would presumably affect only leader TRS-body TRS base-pairing, but not attenuation of nascent minus-strand synthesis (Pasternak *et al.*, 2001). Consequently, the accumulation of attenuated ('anti-leaderless') minus strands would be anticipated. Detection of these molecules, however, is likely to be complicated because (i) they may be subject to rapid degradation, (ii) negative strands are synthesized in small quantities (den Boon *et al.*, 1996; van Marle *et al.*, 1999b), and (iii) the nested set structure of plus- and minus-strand RNAs complicates the design of a specific probe for detection of anti-leaderless minus strands.

An alternative experimental approach is the addition to the system of an exogenous RNA molecule, mimicking a specific transcription template or intermediate. Following its transfection or *in situ* synthesis, the fate and processing of such a molecule could be followed. However, all efforts to engage transfected (plus- or minus-sense) sg RNA molecules in virus-specific transcriptional processes have failed thus far (Makino *et al.*, 1991; Chang *et al.*, 1994), presumably because they lack the proper RNA signals/conformation. The establishment of transcription complexes may require *cis*-acting signals in template and/or transcriptase or, alternatively, transfected sg RNA molecules may be prevented from entering the membrane-bound replication complex (even though transfected DI RNA replicons can be successfully replicated by the viral machinery). On the other hand, Hiscox *et al.* (1995) have shown that coronavirus leaderless minus strands transcribed *in situ* by the vaccinia virus-derived T7 RNA polymerase could be converted into sg plus strands. In the EAV system, analogous experiments failed to yield any sign of amplification of such synthetic transcription intermediates (A. O. Pasternak & E. J. Snijder, unpublished data). In summary, it is obvious that an *in vitro* reconstitution system, in combination with the (forward and reverse) genetics approach, is required to provide the ultimate answers to the question 'How does the system work?' but a reliable system for *in vitro* RNA synthesis has not been developed for any nidovirus thus far.

Micro- and macro-evolutionary aspects of nidovirus transcription

Using reverse genetics, it became possible to engineer mutations that influenced the production of specific, individual sg mRNAs of EAV and – consequently – of the corresponding structural proteins (Pasternak *et al.*, 2000). Virus infectivity was found to strictly correlate with the amounts of different sg mRNA species, suggesting that the structural proteins that they encode are limiting factors during the life cycle of the virus and that the production of these proteins is tightly regulated at the transcriptional level. Analysis of the virus titres of the LA₇U mutant of EAV (Pasternak *et al.*, 2003) also confirmed this notion: the modest decrease in the amount of RNA₇ in this mutant correlated with a drop in virus titre, even though transcription of

all the other major sg mRNA species was enhanced (A. O. Pasternak & E. J. Snijder, unpublished data). Obviously, in order to guarantee the optimal fitness of the virus, the leader and body TRSs must have synergistically evolved so that a tightly regulated, balanced production of every sg mRNA species is ensured. Whereas the leader TRS, being located in an untranslated region, could have evolved without being restricted by protein-coding requirements, the mechanism of the micro-evolution of the body TRSs must have been more complicated, in particular in arteriviruses where, the borders of all genes overlap and, consequently, body TRSs reside in the coding sequence of an upstream gene, a situation that is relatively rare in coronaviruses.

At the 'macro-evolution' level, the recently described absence of a common 5' leader in sg mRNAs of roniviruses and toroviruses is surprising and intriguing, in particular since two mechanisms appear to be united in one and the same system in the case of toroviruses (van Vliet *et al.*, 2002). Rather than being viewed as fundamental, these differences may transpire to be 'variations on a theme', even though they raise many questions about the evolution of nidovirus transcription. They suggest that the development of machinery for discontinuous transcription was a relatively late event in nidovirus evolution, and may in fact have occurred independently in the arterivirus and coronavirus branches after the divergence of these two lineages from a putative common nidovirus ancestor. Alternatively, but conceptually less straightforward, toroviruses and roniviruses may have recently lost the ability to add a common leader sequence to their sg mRNAs. To evaluate these two possibilities, it would be interesting to understand how the use of a common leader sequence influences virus fitness. Assuming coronavirus and arterivirus transcription occurs from anti-leader-containing, sg-length minus-strand RNAs, the incorporation in these templates of regulatory *cis*-acting sequences present in the 3'-proximal domain of the antigenome would allow the use of the same signals for the synthesis of both genome-length and sg plus-strand RNAs. Furthermore, leader addition can be viewed as a strategy to provide all sg mRNAs with a translational enhancer (Tahara *et al.*, 1994) or – in the case of persistent infection – attenuator (Hofmann *et al.*, 1993).

The RNA-strand transfer mechanism used during fusion of sg RNA leader and body segments may require specific protein factors, lacking in nidoviruses that do not employ discontinuous transcription. Interestingly, toroviruses may represent an intermediate stage in the evolution of nidovirus transcription, whatever its 'direction' may have been. Despite the proposed common ancestry of the core enzymes driving nidovirus RNA synthesis, it is likely that future studies will uncover more surprising differences between individual nidovirus subgroups. Thus, the continued functional dissection of the replicative proteins of different nidoviruses will help to clarify some of the outstanding issues concerning the evolution of nidovirus transcription.

Acknowledgements

For pleasant contacts/collaborations and helpful discussions through the years, we would like to thank our past and present colleagues at the Molecular Virology Laboratory, Department of Medical Microbiology of Leiden University Medical Center, and many colleagues from the

nidovirus field. A. O. P. was supported by grant 700-31-020 from the Council for Chemical Sciences of The Netherlands Organization for Scientific Research.

References

- Adkins, S., Siegel, R. W., Sun, J. H. & Kao, C. C. (1997).** Minimal templates directing accurate initiation of subgenomic RNA synthesis in vitro by the brome mosaic virus RNA-dependent RNA polymerase. *RNA* **3**, 634–647.
- Almazan, F., Gonzalez, J. M., Penzes, Z., Izeta, A., Calvo, E., Plana-Duran, J. & Enjuanes, L. (2000).** Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. *Proc Natl Acad Sci U S A* **97**, 5516–5521.
- Almazan, F., Galan, C. & Enjuanes, L. (2004).** The nucleoprotein is required for efficient coronavirus genome replication. *J Virol* **78**, 12683–12688.
- Alonso, S., Izeta, A., Sola, I. & Enjuanes, L. (2002).** Transcription regulatory sequences and mRNA expression levels in the coronavirus transmissible gastroenteritis virus. *J Virol* **76**, 1293–1308.
- An, S. & Makino, S. (1998).** Characterizations of coronavirus *cis*-acting RNA elements and the transcription step affecting its transcription efficiency. *Virology* **243**, 198–207.
- Baker, S. C. & Lai, M. M. C. (1990).** An in vitro system for the leader-primed transcription of coronavirus mRNAs. *EMBO J* **9**, 4173–4179.
- Baric, R. S. & Yount, B. (2000).** Subgenomic negative-strand RNA function during mouse hepatitis virus infection. *J Virol* **74**, 4039–4046.
- Baric, R. S., Stohlman, S. A. & Lai, M. M. C. (1983).** Characterization of replicative intermediate RNA of mouse hepatitis virus: presence of leader RNA sequences on nascent chains. *J Virol* **48**, 633–640.
- Bhardwaj, K., Guarino, L. & Kao, C. C. (2004).** The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. *J Virol* **78**, 12218–12224.
- Brian, D. A. & Spaan, W. J. M. (1997).** Recombination and coronavirus defective interfering RNAs. *Semin Virol* **8**, 101–111.
- Brockway, S. M., Clay, C. T., Lu, X. T. & Denison, M. R. (2003).** Characterization of the expression, intracellular localization, and replication complex association of the putative mouse hepatitis virus RNA-dependent RNA polymerase. *J Virol* **77**, 10515–10527.
- Carpenter, C. D., Oh, J. W., Zhang, C. & Simon, A. E. (1995).** Involvement of a stem-loop structure in the location of junction sites in viral RNA recombination. *J Mol Biol* **245**, 608–622.
- Casais, R., Thiel, V., Siddell, S. G., Cavanagh, D. & Britton, P. (2001).** Reverse genetics system for the avian coronavirus infectious bronchitis virus. *J Virol* **75**, 12359–12369.
- Cascone, P. J., Carpenter, C. D., Li, X. H. & Simon, A. E. (1990).** Recombination between satellite RNAs of turnip crinkle virus. *EMBO J* **9**, 1709–1715.
- Cascone, P. J., Haydar, T. F. & Simon, A. E. (1993).** Sequences and structures required for recombination between virus-associated RNAs. *Science* **260**, 801–805.
- Cavanagh, D. (1997).** Nidovirales: a new order comprising *Coronaviridae* and *Arteriviridae*. *Arch Virol* **142**, 629–633.
- Chang, R. Y., Hofmann, M. A., Sethna, P. B. & Brian, D. A. (1994).** A *cis*-acting function for the coronavirus leader in defective interfering RNA replication. *J Virol* **68**, 8223–8231.

- Chang, R. Y., Krishnan, R. & Brian, D. A. (1996).** The UCUAAAC promoter motif is not required for high-frequency leader recombination in bovine coronavirus defective interfering RNA. *J Virol* **70**, 2720–2729.
- Choi, I. R. & White, K. A. (2002).** An RNA activator of subgenomic mRNA1 transcription in tomato bushy stunt virus. *J Biol Chem* **277**, 3760–3766.
- Choi, I. R., Ostrovsky, M., Zhang, G. & White, K. A. (2001).** Regulatory activity of distal and core RNA elements in tombusvirus subgenomic mRNA2 transcription. *J Biol Chem* **276**, 41761–41768.
- Coley, S. E., Lavi, E., Sawicki, S. G., Fu, L., Schelle, B., Karl, N., Siddell, S. G. & Thiel, V. (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.
- Cowley, J. A., Dimmock, C. M., Spann, K. M. & Walker, P. J. (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.
- Cowley, J. A., Dimmock, C. M. & Walker, P. J. (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.
- Curtis, K. M., Yount, B. & Baric, R. S. (2002).** Heterologous gene expression from transmissible gastroenteritis virus replicon particles. *J Virol* **76**, 1422–1434.
- Curtis, K. M., Yount, B., Sims, A. C. & Baric, R. S. (2004).** Reverse genetic analysis of the transcription regulatory sequence of the coronavirus transmissible gastroenteritis virus. *J Virol* **78**, 6061–6066.
- den Boon, J. A., Snijder, E. J., Chirnside, E. D., de Vries, A. A. F., Horzinek, M. C. & Spaan, W. J. M. (1991).** Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. *J Virol* **65**, 2910–2920.
- den Boon, J. A., Spaan, W. J. M. & Snijder, E. J. (1995).** Equine arteritis virus subgenomic RNA transcription: UV inactivation and translation inhibition studies. *Virology* **213**, 364–372.
- den Boon, J. A., Kleijnen, M. F., Spaan, W. J. M. & Snijder, E. J. (1996).** Equine arteritis virus subgenomic mRNA synthesis: analysis of leader-body junctions and replicative-form RNAs. *J Virol* **70**, 4291–4298.
- de Vries, A. A. F., Chirnside, E. D., Bredenbeek, P. J., Gravesteyn, L. A., Horzinek, M. C. & Spaan, W. J. M. (1990).** All subgenomic mRNAs of equine arteritis virus contain a common leader sequence. *Nucleic Acids Res* **18**, 3241–3247.
- de Vries, A. A. F., Horzinek, M. C., Rottier, P. J. M. & de Groot, R. J. (1997).** The genome organization of the *Nidovirales*: similarities and differences between arteri-, toro-, and coronaviruses. *Semin Virol* **8**, 33–47.
- de Vries, A. A. F., Glaser, A. L., Raamsman, M. J. & Rottier, P. J. M. (2001).** Recombinant equine arteritis virus as an expression vector. *Virology* **284**, 259–276.
- Drosten, C., Gunther, S., Preiser, W. & 23 other authors (2003).** Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med* **348**, 1967–1976.

Egloff, M. P., Ferron, F., Campanacci, V. & 7 other authors (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.

Fischer, F., Stegen, C. F., Koetzner, C. A. & Masters, P. S. (1997). Analysis of a recombinant mouse hepatitis virus expressing a foreign gene reveals a novel aspect of coronavirus transcription. *J Virol* **71**, 5148–5160.

Fouchier, R. A., Hartwig, N. G., Bestebroer, T. M., Niemeyer, B., de Jong, J. C., Simon, J. H. & Osterhaus, A. D. (2004). A previously undescribed coronavirus associated with respiratory disease in humans. *Proc Natl Acad Sci U S A* **101**, 6212–6216.

Fu, K. & Baric, R. S. (1992). Evidence for variable rates of recombination in the MHV genome. *Virology* **189**, 88–102.

Fu, K. & Baric, R. S. (1994). Map locations of mouse hepatitis virus temperature-sensitive mutants: confirmation of variable rates of recombination. *J Virol* **68**, 7458–7466.

Gonzalez, J. M., Gomez-Puertas, P., Cavanagh, D., Gorbalenya, A. E. & Enjuanes, L. (2003). A comparative sequence analysis to revise the current taxonomy of the family *Coronaviridae*. *Arch Virol* **148**, 2207–2235.

Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P. & Blinov, V. M. (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.

Guarino, L. A., Bhardwaj, K., Dong, W., Sun, J., Holzenburg, A. & Kao, C. C. (2005). Mutational analysis of the SARS virus nsp15 endoribonuclease: identification of residues affecting hexamer formation. *J Mol Biol* **353**, 1106–1117.

Guo, J., Wu, T., Anderson, J., Kane, B. F., Johnson, D. G., Gorelick, R. J., Henderson, L. E. & Levin, J. G. (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.

Haasnoot, P. C., Brederode, F. T., Olsthoorn, R. C. & Bol, J. F. (2000). A conserved hairpin structure in *Alfamovirus* and *Bromovirus* subgenomic promoters is required for efficient RNA synthesis in vitro. *RNA* **6**, 708–716.

Hiscox, J. A., Mawditt, K. L., Cavanagh, D. & Britton, P. (1995). Investigation of the control of coronavirus subgenomic mRNA transcription by using T7-generated negative-sense RNA transcripts. *J Virol* **69**, 6219–6227.

Hofmann, M. A., Senanayake, S. D. & Brian, D. A. (1993). A translation-attenuating intraleader open reading frame is selected on coronavirus mRNAs during persistent infection. *Proc Natl Acad Sci U S A* **90**, 11733–11737.

Hsue, B. & Masters, P. S. (1999). Insertion of a new transcriptional unit into the genome of mouse hepatitis virus. *J Virol* **73**, 6128–6135.

Hussain, S., Pan, J., Chen, Y. & 8 other authors (2005). Identification of novel subgenomic RNAs and noncanonical transcription initiation signals of severe acute respiratory syndrome coronavirus. *J Virol* **79**, 5288–5295.

- Ivanov, K. A., Hertzog, T., Rozanov, M., Bayer, S., Thiel, V., Gorbalenya, A. E. & Ziebuhr, J. (2004).** Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc Natl Acad Sci U S A* **101**, 12694–12699.
- Jacobs, L., Spaan, W. J. M., Horzinek, M. C. & van der Zeijst, B. A. (1981).** Synthesis of subgenomic mRNA's of mouse hepatitis virus is initiated independently: evidence from UV transcription mapping. *J Virol* **39**, 401–406.
- Jarvis, T. C. & Kirkegaard, K. (1991).** The polymerase in its labyrinth: mechanisms and implications of RNA recombination. *Trends Genet* **7**, 186–191.
- Jeong, Y. S. & Makino, S. (1994).** Evidence for coronavirus discontinuous transcription. *J Virol* **68**, 2615–2623.
- Jeong, Y. S., Repass, J. F., Kim, Y. N., Hwang, S. M. & Makino, S. (1996).** Coronavirus transcription mediated by sequences flanking the transcription consensus sequence. *Virology* **217**, 311–322.
- Joo, M. & Makino, S. (1992).** Mutagenic analysis of the coronavirus intergenic consensus sequence. *J Virol* **66**, 6330–6337.
- Joo, M. & Makino, S. (1995).** The effect of two closely inserted transcription consensus sequences on coronavirus transcription. *J Virol* **69**, 272–280.
- Keck, J. G., Matsushima, G. K., Makino, S., Fleming, J. O., Vannier, D. M., Stohlman, S. A. & Lai, M. M. C. (1988a).** In vivo RNA-RNA recombination of coronavirus in mouse brain. *J Virol* **62**, 1810–1813.
- Keck, J. G., Soe, L. H., Makino, S., Stohlman, S. A. & Lai, M. M. C. (1988b).** RNA recombination of murine coronaviruses: recombination between fusion-positive mouse hepatitis virus A59 and fusion-negative mouse hepatitis virus 2. *J Virol* **62**, 1989–1998.
- Kim, K. H. & Hemenway, C. L. (1999).** Long-distance RNA–RNA interactions and conserved sequence elements affect potato virus X plus-strand RNA accumulation. *RNA* **5**, 636–645.
- Klovins, J., Berzins, V. & van Duin, J. (1998).** A long-range interaction in Qbeta RNA that bridges the thousand nucleotides between the M-site and the 3' end is required for replication. *RNA* **4**, 948–957.
- Koev, G., Mohan, B. R. & Miller, W. A. (1999).** Primary and secondary structural elements required for synthesis of barley yellow dwarf virus subgenomic RNA1. *J Virol* **73**, 2876–2885.
- Komissarova, N. & Kashlev, M. (1997a).** RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and the RNA. *J Biol Chem* **272**, 15329–15338.
- Komissarova, N. & Kashlev, M. (1997b).** Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded. *Proc Natl Acad Sci U S A* **94**, 1755–1760.
- Konings, D. A., Bredenbeek, P. J., Noten, J. F., Hogeweg, P. & Spaan, W. J. M. (1988).** Differential premature termination of transcription as a proposed mechanism for the regulation of coronavirus gene expression. *Nucleic Acids Res* **16**, 10849–10860.
- Krishnan, R., Chang, R. Y. & Brian, D. A. (1996).** Tandem placement of a coronavirus promoter results in enhanced mRNA synthesis from the downstream-most initiation site. *Virology* **218**, 400–405.

- Ksiazek, T. G., Erdman, D., Goldsmith, C. S. & 23 other authors (2003).** A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med* **348**, 1953–1966.
- Lai, M. M. C. (1998).** Cellular factors in the transcription and replication of viral RNA genomes: a parallel to DNA-dependent RNA transcription. *Virology* **244**, 1–12.
- Lai, M. M. C. & Cavanagh, D. (1997).** The molecular biology of coronaviruses. *Adv Virus Res* **48**, 1–100.
- Lai, M. M. C. & Holmes, K. V. (2001).** *Coronaviridae*. In *Fields Virology*, pp. 1163–1185. Edited by D. M. Knipe & P. M. Howley. Philadelphia, PA: Lippincott, Williams & Wilkins.
- Lai, M. M. C., Patton, C. D. & Stohlman, S. A. (1982a).** Further characterization of mRNAs of mouse hepatitis virus: presence of common 5'-end nucleotides. *J Virol* **41**, 557–565.
- Lai, M. M. C., Patton, C. D. & Stohlman, S. A. (1982b).** Replication of mouse hepatitis virus: negative-stranded RNA and replicative form RNA are of genome length. *J Virol* **44**, 487–492.
- Lai, M. M. C., Patton, C. D., Baric, R. S. & Stohlman, S. A. (1983).** Presence of leader sequences in the mRNA of mouse hepatitis virus. *J Virol* **46**, 1027–1033.
- Lai, M. M. C., Baric, R. S., Brayton, P. R. & Stohlman, S. A. (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.
- Lai, M. M. C., Baric, R. S., Makino, S., Keck, J. G., Egbert, J., Leibowitz, J. L. & Stohlman, S. A. (1985).** Recombination between nonsegmented RNA genomes of murine coronaviruses. *J Virol* **56**, 449–456.
- Liao, C. L. & Lai, M. M. C. (1992).** RNA recombination in a coronavirus: recombination between viral genomic RNA and transfected RNA fragments. *J Virol* **66**, 6117–6124.
- Lin, H. X. & White, K. A. (2004).** A complex network of RNA–RNA interactions controls subgenomic mRNA transcription in a tombusvirus. *EMBO J* **23**, 3365–3374.
- Lin, Y. J., Liao, C. L. & Lai, M. M. C. (1994).** Identification of the *cis*-acting signal for minus-strand RNA synthesis of a murine coronavirus: implications for the role of minus-strand RNA in RNA replication and transcription. *J Virol* **68**, 8131–8140.
- Lin, Y. J., Zhang, X., Wu, R. C. & Lai, M. M. C. (1996).** The 3' untranslated region of coronavirus RNA is required for subgenomic mRNA transcription from a defective interfering RNA. *J Virol* **70**, 7236–7240.
- Lindenbach, B. D., Sgro, J. Y. & Ahlquist, P. (2002).** Long-distance base pairing in flock house virus RNA1 regulates subgenomic RNA3 synthesis and RNA2 replication. *J Virol* **76**, 3905–3919.
- Makino, S., Stohlman, S. A. & Lai, M. M. C. (1986).** Leader sequences of murine coronavirus mRNAs can be freely reassorted: evidence for the role of free leader RNA in transcription. *Proc Natl Acad Sci U S A* **83**, 4204–4208.
- Makino, S., Soe, L. H., Shieh, C. K. & Lai, M. M. C. (1988).** Discontinuous transcription generates heterogeneity at the leader fusion sites of coronavirus mRNAs. *J Virol* **62**, 3870–3873.
- Makino, S., Joo, M. & Makino, J. K. (1991).** A system for study of coronavirus mRNA synthesis: a regulated, expressed subgenomic defective interfering RNA results from intergenic site insertion. *J Virol* **65**, 6031–6041.

- Masters, P. S., Koetzner, C. A., Kerr, C. A. & Heo, Y. (1994).** Optimization of targeted RNA recombination and mapping of a novel nucleocapsid gene mutation in the coronavirus mouse hepatitis virus. *J Virol* **68**, 328–337.
- Meulenberg, J. J. M., Bos-de Ruijter, J. N. A., van de Graaf, R., Wensvoort, G. & Moormann, R. J. M. (1998).** Infectious transcripts from cloned genome-length cDNA of porcine reproductive respiratory syndrome virus. *J Virol* **72**, 380–387.
- Miller, W. A. & Koev, G. (2000).** Synthesis of subgenomic RNAs by positive-strand RNA viruses. *Virology* **273**, 1–8.
- Miller, W. A., Dreher, T. W. & Hall, T. C. (1985).** Synthesis of brome mosaic virus subgenomic RNA in vitro by internal initiation on (-)-sense genomic RNA. *Nature* **313**, 68–70.
- Molenkamp, R., Greve, S., Spaan, W. J. M. & Snijder, E. J. (2000a).** Efficient homologous RNA recombination and requirement for an open reading frame during replication of equine arteritis virus defective interfering RNAs. *J Virol* **74**, 9062–9070.
- Molenkamp, R., van Tol, H., Rozier, B. C., van der Meer, Y., Spaan, W. J. M. & Snijder, E. J. (2000b).** The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription. *J Gen Virol* **81**, 2491–2496.
- Nagy, P. D. & Bujarski, J. J. (1993).** Targeting the site of RNA–RNA recombination in brome mosaic virus with antisense sequences. *Proc Natl Acad Sci U S A* **90**, 6390–6394.
- Nagy, P. D. & Bujarski, J. J. (1995).** Efficient system of homologous RNA recombination in brome mosaic virus: sequence and structure requirements and accuracy of crossovers. *J Virol* **69**, 131–140.
- Nagy, P. D. & Bujarski, J. J. (1996).** Homologous RNA recombination in brome mosaic virus: AU-rich sequences decrease the accuracy of crossovers. *J Virol* **70**, 415–426.
- Nagy, P. D. & Simon, A. E. (1997).** New insights into the mechanisms of RNA recombination. *Virology* **235**, 1–9.
- Nagy, P. D. & Simon, A. E. (1998a).** In vitro characterization of late steps of RNA recombination in turnip crinkle virus. I. Role of motif1-hairpin structure. *Virology* **249**, 379–392.
- Nagy, P. D. & Simon, A. E. (1998b).** In vitro characterization of late steps of RNA recombination in turnip crinkle virus. II. The role of the priming stem and flanking sequences. *Virology* **249**, 393–405.
- Nagy, P. D., Dziafott, A., Ahlquist, P. & Bujarski, J. J. (1995).** Mutations in the helicase-like domain of protein 1a alter the sites of RNA-RNA recombination in brome mosaic virus. *J Virol* **69**, 2547–2556.
- Nagy, P. D., Zhang, C. & Simon, A. E. (1998).** Dissecting RNA recombination in vitro: role of RNA sequences and the viral replicase. *EMBO J* **17**, 2392–2403.
- Nagy, P. D., Pogany, J. & Simon, A. E. (1999).** RNA elements required for RNA recombination function as replication enhancers *in vitro* and *in vivo* in a plus-strand RNA virus. *EMBO J* **18**, 5653–5665.
- Nelson, G. W., Stohlman, S. A. & Tahara, S. M. (2000).** High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA. *J Gen Virol* **81**, 181–188.
- Nudler, E., Avetissova, E., Markovtsov, V. & Goldfarb, A. (1996).** Transcription processivity: protein-DNA interactions holding together the elongation complex. *Science* **273**, 211–217.

Nudler, E., Mustaev, A., Lukhtanov, E. & Goldfarb, A. (1997). The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* **89**, 33–41.

Ozdarendeli, A., Ku, S., Rochat, S., Williams, G. D., Senanayake, S. D. & Brian, D. A. (2001). Downstream sequences influence the choice between a naturally occurring noncanonical and closely positioned upstream canonical heptameric fusion motif during bovine coronavirus subgenomic mRNA synthesis. *J Virol* **75**, 7362–7374.

Pasternak, A. O. (2003). *Nidovirus transcription-regulating sequences*. PhD thesis. Leiden University.

Pasternak, A. O., Gulyaev, A. P., Spaan, W. J. M. & Snijder, E. J. (2000). Genetic manipulation of arterivirus alternative mRNA leader-body junction sites reveals tight regulation of structural protein expression. *J Virol* **74**, 11642–11653.

Pasternak, A. O., van den Born, E., Spaan, W. J. M. & Snijder, E. J. (2001). Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. *EMBO J* **20**, 7220–7228.

Pasternak, A. O., van den Born, E., Spaan, W. J. M. & Snijder, E. J. (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.

Pasternak, A. O., Spaan, W. J. M. & Snijder, E. J. (2004). Regulation of relative abundance of arterivirus subgenomic mRNAs. *J Virol* **78**, 8102–8113.

Peiris, J. S., Lai, S. T., Poon, L. L. & 13 other authors (2003). Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* **361**, 1319–1325.

Peliska, J. A. & Benkovic, S. J. (1992). Mechanism of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase. *Science* **258**, 1112–1118.

Peliska, J. A. & Benkovic, S. J. (1994). Fidelity of in vitro DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase. *Biochemistry* **33**, 3890–3895.

Posthuma, C. C., Nedialkova, D. D., Zevenhoven-Dobbe, J. C., Blokhuis, J. H., Gorbalenya, A. E. & Snijder, E. J. (2006). Site-directed mutagenesis of the nidovirus replicative endoribonuclease NendoU exerts pleiotropic effects on the arterivirus life cycle. *J Virol* **80**, 1653–1661.

Reeder, R. H. & Lang, W. H. (1997). Terminating transcription in eukaryotes: lessons learned from RNA polymerase I. *Trends Biochem Sci* **22**, 473–477.

Satyanarayana, T., Gowda, S., Ayllon, M. A., Albiach-Marti, M. R., Rabindran, S. & Dawson, W. O. (2002). The p23 protein of *Citrus tristeza virus* controls asymmetrical RNA accumulation. *J Virol* **76**, 473–483.

Sawicki, S. G. & Sawicki, D. L. (1990). Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. *J Virol* **64**, 1050–1056.

Sawicki, S. G. & Sawicki, D. L. (1995). Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. *Adv Exp Med Biol* **380**, 499–506.

Sawicki, S. G. & Sawicki, D. L. (2005). Coronavirus transcription: a perspective. *Curr Top Microbiol Immunol* **287**, 31–55.

Sawicki, D. L., Wang, T. & Sawicki, S. G. (2001). The RNA structures engaged in replication and transcription of the A59 strain of mouse hepatitis virus. *J Gen Virol* **82**, 385–396.

- Sawicki, S. G., Sawicki, D. L., Younker, D., Meyer, Y., Thiel, V., Stokes, H. & Siddell, S. G. (2005).** Functional and genetic analysis of coronavirus replicase-transcriptase proteins. *PLoS Pathog* **4**, 310–322.
- Schaad, M. C., Stohlman, S. A., Egbert, J., Lum, K., Fu, K., Wei, T. Jr & Baric, R. S. (1990).** Genetics of mouse hepatitis virus transcription: identification of cistrons which may function in positive and negative strand RNA synthesis. *Virology* **177**, 634–645.
- Schelle, B., Karl, N., Ludewig, B., Siddell, S. G. & Thiel, V. (2005).** Selective replication of coronavirus genomes that express nucleocapsid protein. *J Virol* **79**, 6620–6630.
- Sengupta, D. J., Wickens, M. & Fields, S. (1999).** Identification of RNAs that bind to a specific protein using the yeast three-hybrid system. *RNA* **5**, 596–601.
- Sethna, P. B., Hung, S. L. & Brian, D. A. (1989).** Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. *Proc Natl Acad Sci U S A* **86**, 5626–5630.
- Seybert, A., van Dinten, L. C., Snijder, E. J. & Ziebuhr, J. (2000).** Biochemical characterization of the equine arteritis virus helicase suggests a close functional relationship between arterivirus and coronavirus helicases. *J Virol* **74**, 9586–9593.
- Seybert, A., Posthuma, C. C., van Dinten, L. C., Snijder, E. J., Gorbalenya, A. E. & Ziebuhr, J. (2005).** A complex zinc finger controls the enzymatic activities of nidovirus helicases. *J Virol* **79**, 696–704.
- Siddell, S. G., Ziebuhr, J. & Snijder, E. J. (2005).** In *Topley and Wilson's Microbiology and Microbial Infections; Virology Volume*, 10th edn, pp. 823–856. Edited by B. W. Mahy & V. ter Meulen. London: Hodder Arnold.
- Siegel, R. W., Adkins, S. & Kao, C. C. (1997).** Sequence-specific recognition of a subgenomic RNA promoter by a viral RNA polymerase. *Proc Natl Acad Sci U S A* **94**, 11238–11243.
- Siegel, R. W., Bellon, L., Beigelman, L. & Kao, C. C. (1998).** Moieties in an RNA promoter specifically recognized by a viral RNA-dependent RNA polymerase. *Proc Natl Acad Sci U S A* **95**, 11613–11618.
- Siegel, R. W., Bellon, L., Beigelman, L. & Kao, C. C. (1999).** Use of DNA, RNA, and chimeric templates by a viral RNA-dependent RNA polymerase: evolutionary implications for the transition from the RNA to the DNA world. *J Virol* **73**, 6424–6429.
- Sit, T. L., Vaewhongs, A. A. & Lommel, S. A. (1998).** RNA-mediated trans-activation of transcription from a viral RNA. *Science* **281**, 829–832.
- Sivakumaran, K. & Kao, C. C. (1999).** Initiation of genomic plus-strand RNA synthesis from DNA and RNA templates by a viral RNA-dependent RNA polymerase. *J Virol* **73**, 6415–6423.
- Smits, S. L., van Vliet, A. L., Segeren, K., el Azzouzi, H., van Essen, M. & de Groot, R. J. (2005).** Torovirus non-discontinuous transcription: mutational analysis of a subgenomic mRNA promoter. *J Virol* **79**, 8275–8281.
- Snijder, E. J. & Horzinek, M. C. (1993).** Toroviruses: replication, evolution and comparison with other members of the coronavirus-like superfamily. *J Gen Virol* **74**, 2305–2316.
- Snijder, E. J. & Meulenberg, J. J. M. (1998).** The molecular biology of arteriviruses. *J Gen Virol* **79**, 961–979.

Snijder, E. J., den Boon, J. A., Bredenbeek, P. J., Horzinek, M. C., Rijnbrand, R. & Spaan, W. J. M. (1990a). The carboxyl-terminal part of the putative Berne virus polymerase is expressed by ribosomal frameshifting and contains sequence motifs which indicate that toro- and coronaviruses are evolutionarily related. *Nucleic Acids Res* **18**, 4535–4542.

Snijder, E. J., Horzinek, M. C. & Spaan, W. J. M. (1990b). A 3'-coterminal nested set of independently transcribed mRNAs is generated during Berne virus replication. *J Virol* **64**, 331–338.

Snijder, E. J., Bredenbeek, P. J., Dobbe, J. C. & 7 other authors (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.

Snijder, E. J., Siddell, S. G. & Gorbalenya, A. E. (2005). The order Nidovirales. In *Topley and Wilson's Microbiology and Microbial Infections; Virology Volume*, pp. 390–404. Edited by B. W. Mahy & V. ter Meulen. London: Hodder Arnold.

Sola, I., Alonso, S., Zuniga, S., Balasch, M., Plana-Duran, J. & Enjuanes, L. (2003). Engineering the transmissible gastroenteritis virus genome as an expression vector inducing lactogenic immunity. *J Virol* **77**, 4357–4369.

Sola, I., Moreno, J. L., Zuniga, S., Alonso, S. & Enjuanes, L. (2005). Role of nucleotides immediately flanking the transcription-regulating sequence core in coronavirus subgenomic mRNA synthesis. *J Virol* **79**, 2506–2516.

Spaan, W. J. M., Rottier, P. J. M., Horzinek, M. C. & van der Zeijst, B. A. M. (1982). Sequence relationships between the genome and the intracellular RNA species 1, 3, 6, and 7 of mouse hepatitis virus strain A59. *J Virol* **42**, 432–439.

Spaan, W. J. M., Delius, H., Skinner, M., Armstrong, J., Rottier, P. J. M., Smeekens, S., van der Zeijst, B. A. M. & Siddell, S. G. (1983). Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. *EMBO J* **2**, 1839–1844.

Stawicki, S. S. & Kao, C. C. (1999). Spatial perturbations within an RNA promoter specifically recognized by a viral RNA-dependent RNA polymerase (RdRp) reveal that RdRp can adjust its promoter binding sites. *J Virol* **73**, 198–204.

Stern, D. F. & Sefton, B. M. (1982). Synthesis of coronavirus mRNAs: kinetics of inactivation of infectious bronchitis virus RNA synthesis by UV light. *J Virol* **42**, 755–759.

Stirrup, K., Shaw, K., Evans, S., Dalton, K., Cavanagh, D. & Britton, P. (2000). Leader switching occurs during the rescue of defective RNAs by heterologous strains of the coronavirus infectious bronchitis virus. *J Gen Virol* **81**, 791–801.

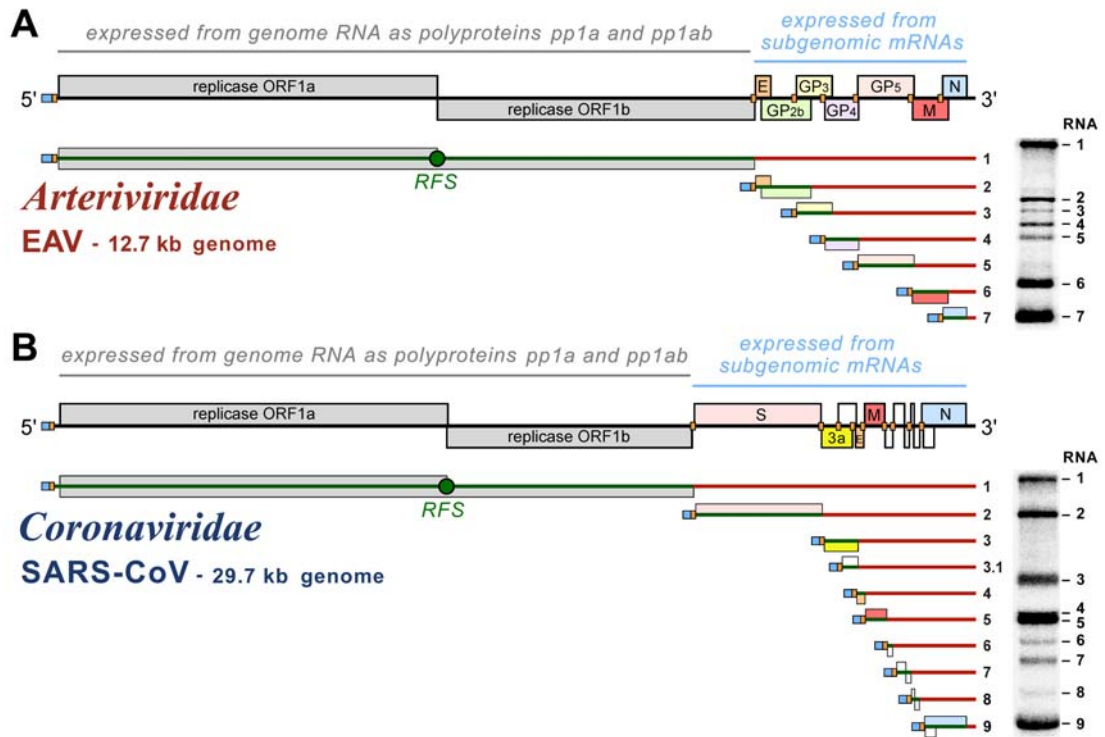
Sturman, L. S., Eastwood, C., Frana, M. F., Duchala, C., Baker, F., Ricard, C. S., Sawicki, S. G. & Holmes, K. V. (1987). Temperature-sensitive mutants of MHV-A59. *Adv Exp Med Biol* **218**, 159–168.

Sutton, G., Fry, E., Carter, L. & 14 other authors (2004). The nsp9 replicase protein of SARS-coronavirus, structure and functional insights. *Structure* **12**, 341–353.

Tahara, S. M., Dietlin, T. A., Bergmann, C. C., Nelson, G. W., Kyuwa, S., Anthony, R. P. & Stohlman, S. A. (1994). Coronavirus translational regulation: leader affects mRNA efficiency. *Virology* **202**, 621–630.

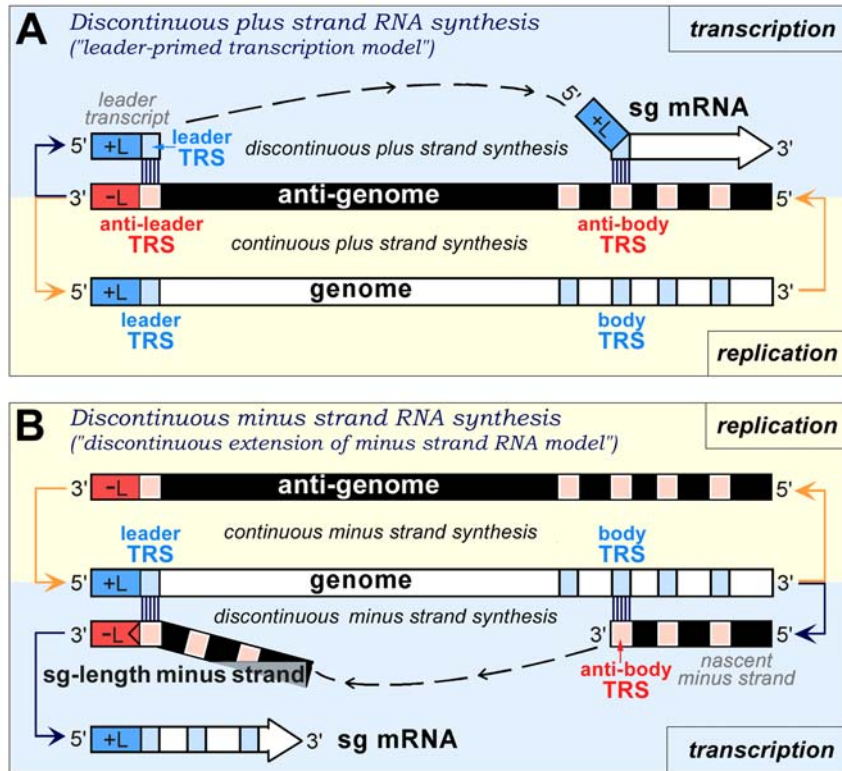
- Thiel, V., Herold, J., Schelle, B. & Siddell, S. G. (2001).** Infectious RNA transcribed in vitro from a cDNA copy of the human coronavirus genome cloned in vaccinia virus. *J Gen Virol* **82**, 1273–1281.
- Tijms, M. A. (2004).** *Nsp1, a multifunctional regulator of the arterivirus life cycle*. PhD thesis. Leiden University.
- Tijms, M. A., van Dinten, L. C., Gorbalenya, A. E. & Snijder, E. J. (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.
- van den Born, E., Gultyaev, A. P. & Snijder, E. J. (2004).** Secondary structure and function of the 5'-proximal region of the equine arteritis virus RNA genome. *RNA* **10**, 424–437.
- van den Born, E., Posthuma, C. C., Gultyaev, A. P. & Snijder, E. J. (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.
- van der Hoek, L., Pyrc, K., Jebbink, M. F. & 7 other authors (2004).** Identification of a new human coronavirus. *Nat Med* **10**, 368–373.
- van der Most, R. G. & Spaan, W. J. M. (1995).** Coronavirus replication, transcription and RNA recombination. In *The Coronaviridae*, pp. 11–31. Edited by S. G. Siddell. Plenum.
- van der Most, R. G., de Groot, R. J. & Spaan, W. J. M. (1994).** Subgenomic RNA synthesis directed by a synthetic defective interfering RNA of mouse hepatitis virus: a study of coronavirus transcription initiation. *J Virol* **68**, 3656–3666.
- van Dinten, L. C., den Boon, J. A., Wassenaar, A. L., Spaan, W. J. M. & Snijder, E. J. (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.
- van Dinten, L. C., van Tol, H., Gorbalenya, A. E. & Snijder, E. J. (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.
- van Marle, G., Luytjes, W., van der Most, R. G., van der Straaten, T. & Spaan, W. J. M. (1995).** Regulation of coronavirus mRNA transcription. *J Virol* **69**, 7851–7856.
- van Marle, G., Dobbe, J. C., Gultyaev, A. P., Luytjes, W., Spaan, W. J. M. & Snijder, E. J. (1999a).** 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.
- van Marle, G., van Dinten, L. C., Spaan, W. J. M., Luytjes, W. & Snijder, E. J. (1999b).** Characterization of an equine arteritis virus replicase mutant defective in subgenomic mRNA synthesis. *J Virol* **73**, 5274–5281.
- van Vliet, A. L., Smits, S. L., Rottier, P. J. M. & de Groot, R. J. (2002).** Discontinuous and non-discontinuous subgenomic RNA transcription in a nidovirus. *EMBO J* **21**, 6571–6580.
- von Grothuss, M., Wyrwicz, L. S. & Rychlewski, L. (2003).** mRNA cap-1 methyltransferase in the SARS genome. *Cell* **113**, 701–702.
- Wang, J., Carpenter, C. D. & Simon, A. E. (1999).** Minimal sequence and structural requirements of a subgenomic RNA promoter for turnip crinkle virus. *Virology* **253**, 327–336.
- White, K. A. (2002).** The premature termination model: a possible third mechanism for subgenomic mRNA transcription in (+)-strand RNA viruses. *Virology* **304**, 147–154.

- White, K. A. & Morris, T. J. (1995).** RNA determinants of junction site selection in RNA virus recombinants and defective interfering RNAs. *RNA* **1**, 1029–1040.
- Wilson, K. S. & von Hippel, P. H. (1995).** Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proc Natl Acad Sci U S A* **92**, 8793–8797.
- Woo, P. C., Lau, S. K., Chu, C. M. & 12 other authors (2005).** Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J Virol* **79**, 884–895.
- Wu, W., Blumberg, B. M., Fay, P. J. & Bambara, R. A. (1995).** Strand transfer mediated by human immunodeficiency virus reverse transcriptase *in vitro* is promoted by pausing and results in misincorporation. *J Biol Chem* **270**, 325–332.
- Yan, L., Velikanov, M., Flook, P., Zheng, W., Szalma, S. & Kahn, S. (2003).** Assessment of putative protein targets derived from the SARS genome. *FEBS Lett* **554**, 257–263.
- Yokomori, K., Banner, L. R. & Lai, M. M. (1992).** Coronavirus mRNA transcription: UV light transcriptional mapping studies suggest an early requirement for a genomic-length template. *J Virol* **66**, 4671–4678.
- Yount, B., Denison, M. R., Weiss, S. R. & Baric, R. S. (2002).** Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. *J Virol* **76**, 11065–11078.
- Yount, B., Curtis, K. M., Fritz, E. A., Hensley, L. E., Jahrling, P. B., Prentice, E., Denison, M. R., Geisbert, T. W. & Baric, R. S. (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.
- Zhai, Y., Sun, F., Li, X., Pang, H., Xu, X., Bartlam, M. & Rao, Z. (2005).** Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer. *Nat Struct Mol Biol* **12**, 980–986.
- Zhang, X. & Lai, M. M. C. (1994).** Unusual heterogeneity of leader-mRNA fusion in a murine coronavirus: implications for the mechanism of RNA transcription and recombination. *J Virol* **68**, 6626–6633.
- Zhang, X., Liao, C. L. & Lai, M. M. C. (1994).** Coronavirus leader RNA regulates and initiates subgenomic mRNA transcription both in *trans* and in *cis*. *J Virol* **68**, 4738–4746.
- Zhang, G., Slowinski, V. & White, K. A. (1999).** Subgenomic mRNA regulation by a distal RNA element in a (+)-strand RNA virus. *RNA* **5**, 550–561.
- Ziebuhr, J., Snijder, E. J. & Gorbalenya, A. E. (2000).** Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J Gen Virol* **81**, 853–879.
- Zuniga, S., Sola, I., Alonso, S. & Enjuanes, L. (2004).** Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *J Virol* **78**, 980–994.



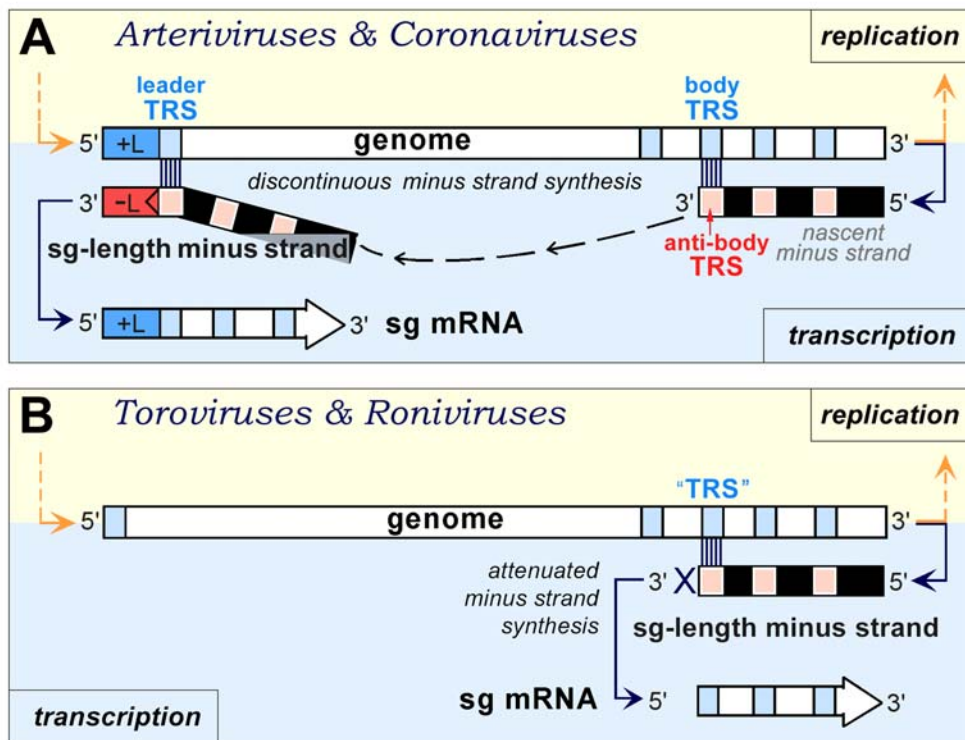
JGV 81611 - Pasternak et al. - Fig. 1

Fig. 1. Nidovirus genomic structure and genome expression strategy, using the arterivirus EAV (a) and the coronavirus SARS-CoV (b) as examples. The open reading frames in the viral genomes are indicated at the top of each schematic. Note that the arterivirus and coronavirus genomes are drawn to different scales. The names of the replicase gene and structural protein genes are given (references to the nomenclature of genes and proteins can be found in Snijder *et al.*, 2005). Below the genome organization, the structural relationships of the genome and sg mRNAs are depicted. The leader sequence and TRSs found at the 5' end of all EAV and SARS-CoV mRNAs are indicated as blue and orange boxes, respectively. The ribosomal frameshifting element (RFS) found in the genome-length mRNA1 is indicated and the translated region of each mRNA is highlighted by a green line, whereas translationally silent regions are indicated by a red line. Only the translated open reading frames are indicated for each mRNA. The right-hand panels show a typical pattern of EAV and SARS-CoV mRNAs isolated from infected cells, visualized by hybridization to a probe complementary to the 3' end of the genome and therefore recognizing all viral mRNA species. SARS-CoV RNA3.1 is a recently described minor sg mRNA, which is not visible in the hybridization analysis (Hussain *et al.*, 2005).



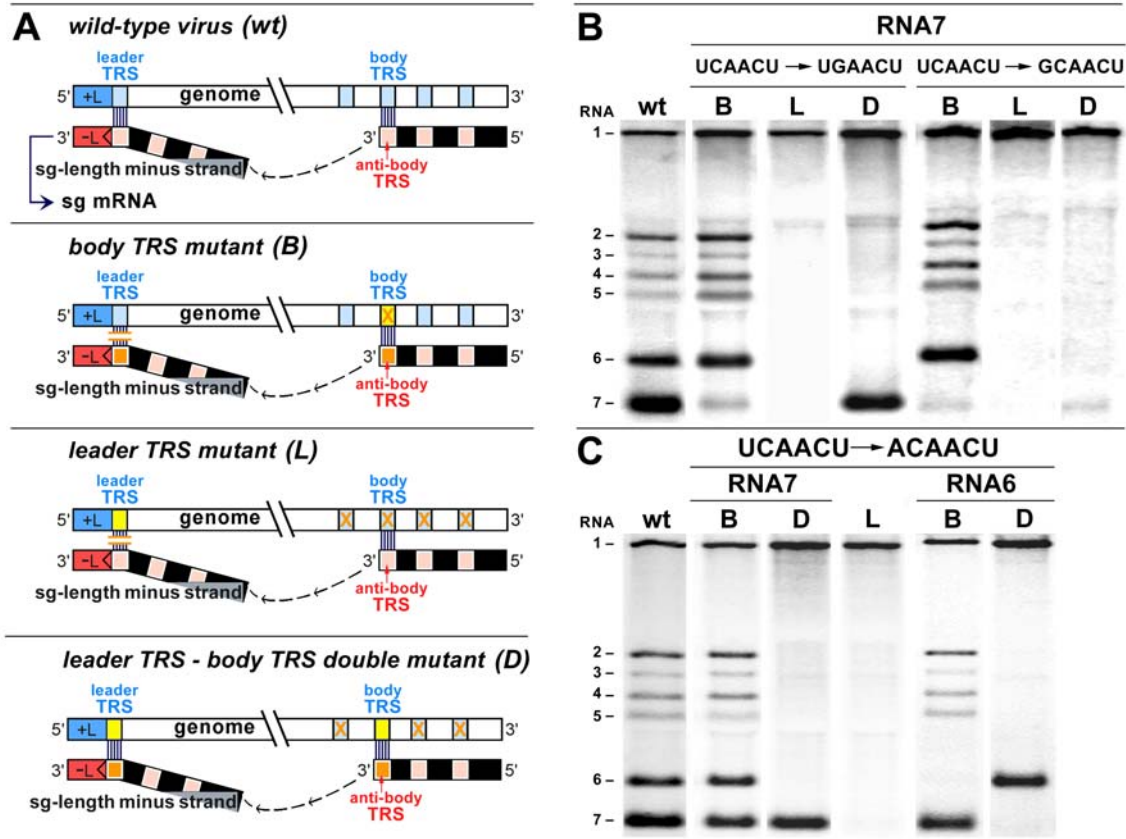
JGV 81611 - Pasternak et al. - Fig. 2

Fig. 2. Opposing models for transcription in arteriviruses and coronaviruses (using a hypothetical virus that produces four sg mRNAs). Both replication (yellow background) and transcription (blue background) processes are depicted. (a) The ‘leader-primed transcription model’ (Baric *et al.*, 1983; Spaan *et al.*, 1983; Lai *et al.*, 1984) which proposes that plus-strand RNA synthesis from the antigenome template is either continuous (producing new genome RNA) or discontinuous (yielding sg mRNAs). Following the transcription of a leader primer (+L) from the 3’ end of the antigenome, the leader TRS in this primer would base pair to an anti-body TRS in the antigenome and be extended to produce a sg mRNA. (b) Model based on discontinuous extension of minus-strand RNA synthesis (Sawicki & Sawicki, 1995) which proposes that minus-strand RNA can be either continuous (producing the antigenome) or discontinuous (yielding sg-length minus strands). The body TRSs in the genome would act as attenuation signals for minus-strand RNA synthesis, after which the nascent minus strand, having an anti-body TRS at its 3’ end, would be redirected to the 5’-proximal region of the template, guided by a base-pairing interaction with the leader TRS. Following the addition of the anti-leader (–L) to the nascent minus strands, the sg-length minus strands would serve as templates for transcription. Adapted from Snijder *et al.* (2005).



JGV 81611 - Pasternak et al. - Fig. 3

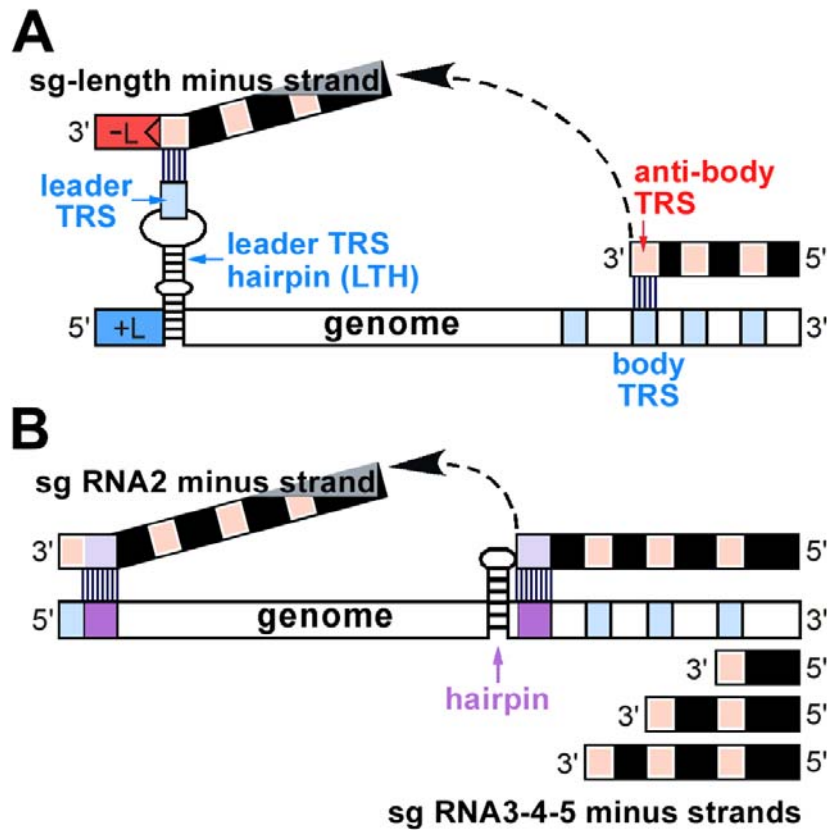
Fig. 3. Models for transcription from minus strand sg-length templates in different nidoviruses (using hypothetical viruses that produce four sg mRNAs). (a) Model for arteriviruses and coronaviruses, which have a common 5' leader sequence on all viral mRNAs. Discontinuous extension of minus-strand RNA synthesis (see also Fig. 2b) has been proposed as the mechanism to produce sg-length minus-strand templates for transcription. The replicase/transcriptase is proposed to be attenuated at one of the body TRSs in the 3'-proximal part of the genome, after which the nascent minus-strand would be extended with the anti-leader (-L) sequence in a process of discontinuous RNA synthesis. Next, the completed sg-length minus strands would serve as templates for transcription. (b) Model for production of sg mRNAs of roniviruses and all but the largest sg mRNAs (see Fig. 5b) of toroviruses, which do not contain a common 5' leader sequence (Cowley *et al.*, 2002; van Vliet *et al.*, 2002). Conserved ('TRS-like') sequences, found upstream of each of the genes in the 3'-proximal part of the genome, are thought to direct a variant of the mechanism explained in panel (a). Attenuation of minus-strand RNA synthesis would occur, but not be followed by the discontinuous step required for addition of the anti-leader sequence (as in panel a). Instead, the attenuated sg-length minus-strand RNAs would serve directly as template for transcription. Adapted from Snijder *et al.* (2005).



JGV 81611 - Pasternak et al. - Fig. 4

Fig. 4. Examples of mutagenesis of EAV TRSs by using reverse genetics (Pasternak *et al.*, 2001). (a) Schematic explanation of the effects of TRS mutagenesis on transcription in arteriviruses and coronaviruses, using a hypothetical virus that produces four sg mRNAs. Using an infectious cDNA clone, point mutations can be introduced into leader TRS and body TRSs. Body TRS mutations (lanes B in panel b) will affect the transcription of a single sg mRNA (EAV RNA 7 in panel b), by disrupting the base-pairing interaction with the leader TRS. Leader TRS mutations (lanes L) will affect the transcription of all sg mRNAs. Upon introduction of the same mutation in leader TRS and a body TRS (double mutants, D), the transcription of a specific sg mRNA can be restored (see panel b). Examples of these three types of mutant constructs are shown in this schematic. Mutant TRSs are depicted in yellow and their complements in orange. Body TRSs that do not yield an sg mRNA in a given mutant are marked with an orange X. (b) Example from the mutagenesis study targeting the 5'-UCAACU-3' consensus sequence of the EAV leader TRS and RNA7 body TRS (Pasternak *et al.*, 2001). Two types of body TRS mutations were identified. The first type (exemplified here by the position 2 mutant 5'-UCAACU-3' → 5'-UGAACU-3' mutation) allowed restoration of RNA7 transcription in the double mutant (lane D). The second type (exemplified by position 1 mutant 5'-UCAACU-3' → 5'-GCAACU-3') did not show such complementation, suggesting an additional function for the body TRS (see

text for details). wt, Wild-type control virus. (c) Example of differential effects of the same mutation at position 1 of the TRS (U to A) in different EAV body TRSs. Upon introduction of the mutation into the RNA7 body TRS, RNA7 transcription (measured as the relative amount of sgRNA7 to gRNA1, to correct for transfection efficiency) was reduced to about 40 %. Introduction of the same mutation into leader TRS (lane D) did not complement this defect. However, in the context of the RNA6 body TRS the same mutation reduced RNA6 transcription to about 1 % and showed restoration to near wild-type levels in the double mutant (D. D. Nedialkova, A. O. Pasternak & E. J. Snijder, unpublished data).



JGV 81611 - Pasternak et al. - Fig. 5

Fig. 5. Examples of RNA structures involved in nidovirus transcription. (a) In the arterivirus EAV, the leader TRS is thought to reside in a stable leader TRS hairpin (LTH) structure, which may facilitate the base-pairing interaction with the anti-body TRS at the 3' end of the incoming nascent minus strand (see also Fig. 2b), a key step in the production of sg-length minus strands. The existence and importance of the LTH was supported by the results from phylogenetic analysis, structure probing and site-directed mutagenesis studies (van den Born *et al.*, 2004, 2005). Similar structures can be predicted for other arteriviruses and most coronaviruses (Chang *et al.*, 1996; van den Born *et al.*, 2004). (b) Model for the transcription of sg mRNA2 of toroviruses as proposed by van Vliet *et al.* (2002). In contrast to sg mRNAs3–5 (see Fig. 3b), sg mRNA2 contains a small 5' leader sequence that is derived from the 5' end of the genome. It was proposed that minus-strand RNA synthesis is attenuated by a hairpin structure in the genome, yielding a nascent strand with a 3'-terminal sequence (depicted in purple) that is highly complementary to a sequence in the genomic 5' end. Following a base-pairing interaction similar to that proposed for coronaviruses and arteriviruses (Fig. 2), the minus strand may be provided with the 5'-terminal sequence of the genome and thus become a template for the transcription of sg mRNA2.