

Published online ahead of print on 1 March 2006 as DOI 10.1099/vir.0.81786-0

REVIEW

Unravelling the complexities of respiratory syncytial virus RNA synthesis

Vanessa M. Cowton, David R. McGivern and Rachel Fearn

Division of Pathology and Neuroscience, University of Dundee Medical School, Dundee
DD1 9SY, UK

Correspondence

Rachel Fearn

r.fearn@dundee.ac.uk

***Human respiratory syncytial virus (RSV)* is the leading cause of paediatric respiratory disease and is the focus of antiviral- and vaccine-development programmes. These goals have been aided by an understanding of the virus genome architecture and the mechanisms by which it is expressed and replicated. RSV is a member of the order *Mononegavirales* and, as such, has a genome consisting of a single strand of negative-sense RNA. At first glance, transcription and genome replication appear straightforward, requiring self-contained promoter regions at the 3' ends of the genome and antigenome RNAs, short *cis*-acting elements flanking each of the genes and one polymerase. However, from these minimal elements, the virus is able to generate an array of capped, methylated and polyadenylated mRNAs and encapsidated antigenome and genome RNAs, all in the appropriate ratios to facilitate virus replication. The apparent simplicity of genome expression and replication is a consequence of considerable complexity in the polymerase structure and its cognate *cis*-acting sequences; here, our understanding of mechanisms by which the RSV polymerase proteins interact with signals in the RNA template to produce different RNA products are reviewed.**

Background and scope of the review

The World Health Organization estimates that *Human respiratory syncytial virus* (RSV) is responsible for 64 million infections and 160 000 deaths per annum. Its victims are mostly young infants, but it is increasingly recognized as a significant cause of disease in the elderly population and can often be fatal for patients with compromised immune systems (Collins *et al.*, 2001). RSV is a member of the subfamily *Pneumovirinae* in the family *Paramyxoviridae*, order *Mononegvirales*, i.e. the non-segmented, negative-strand RNA viruses. This order includes several exotic pathogens, such as the Ebola and Nipah viruses, and other, more familiar ones, such as the parainfluenza, measles and mumps viruses. Most of our initial understanding of mononegavirus transcription and genome replication stemmed from studies with the paramyxovirus *Sendai virus* (SeV) and the rhabdovirus *Vesicular stomatitis virus* (VSV). These viruses can be grown to very high titres, which has facilitated analysis of their RNA-synthesis mechanisms using biochemical techniques. These studies still provide a blueprint for mononegavirus RNA synthesis; however, sequence analysis of different virus genomes and advances in reverse-genetics techniques have opened up this field. Recent investigations have revealed that mononegaviruses differ in the layout of their promoter and gene-junction regions, template structure and polymerase composition. Therefore, although the overall strategy of transcription and replication is similar for all non-segmented, negative-strand RNA viruses, it is possible that the molecular mechanisms that individual viruses use to achieve this differ. Given that RSV is an important human pathogen, it is worth consideration in its own right and, in this review, we aim to examine the data regarding RSV specifically, highlighting where appropriate the similarities and differences with other non-segmented, negative-strand RNA viruses, which might help to identify common themes across the order. In addition, we draw comparisons with well-characterized cellular-transcription models, which provide paradigms for understanding RNA synthesis.

The RSV replication cycle

In the RSV particle, the genome is contained within a helical nucleocapsid, which is associated with virus polymerase proteins and surrounded by matrix protein and an envelope containing viral glycoproteins. The nucleocapsid and polymerase are delivered to the host-cell cytoplasm by direct fusion of the virus particle with the plasma membrane and it is in the cytoplasm that the genome is transcribed and replicated. Transcription results in synthesis of ten monocistronic, capped, methylated and polyadenylated mRNAs, which are translated by the host-cell machinery. Replication involves generation of a complete, positive-sense RNA complement of the genome, the antigenome, which in turn acts as a template for genome synthesis. The genomes are incorporated into nucleocapsids as they are synthesized and can presumably recycle through the RNA-synthesis pathways, or are transported to the plasma

membrane for assembly into virus particles [for a detailed review of the RSV replication cycle, see Collins *et al.* (2001)].

In infected cells, the virus polymerase proteins are detected almost exclusively as large, cytoplasmic inclusions (García *et al.*, 1993), which are presumed to be accumulations of nucleocapsids functioning as RNA-synthesis 'factories'. There is evidence that nucleocapsids are associated with lipid-raft structures (McDonald *et al.*, 2004; Brown *et al.*, 2005) and the cytoskeleton-associated proteins actin and profilin are important for RSV RNA synthesis (García-Barreno *et al.*, 1988; Barik, 1992; Burke *et al.*, 1998, 2000; Ulloa *et al.*, 1998; Kallewaard *et al.*, 2005). These data suggest that RSV RNA-synthesis machinery is integrated intimately with host-cell components and that, when RSV nucleocapsids enter the cell, they co-opt lipid rafts and cytoskeleton proteins to form a structure on which RNA synthesis can occur.

Overview of the RSV transcription and replication strategy

The structure of the RSV genome is shown schematically in Fig. 1. The ten viral genes are arranged sequentially and each is flanked by conserved gene-start (GS) and gene-end (GE) sequences, which control the polymerase during transcription (Collins *et al.*, 1986, 1987; Kuo *et al.*, 1996a, 1997; Fearn & Collins, 1999a). At the 3' and 5' ends of the genome are short extragenic regions: a 44 nt leader (Le) and a 155 nt trailer (Tr), respectively (Collins *et al.*, 1991; Mink *et al.*, 1991). There is just one promoter for transcription and this lies at the 3' end of the genome and includes sequences in the Le (Dickens *et al.*, 1984; Collins *et al.*, 1991). The polymerase initiates RNA synthesis at this promoter and then progresses along the length of the genome. When it reaches a GE signal, the polymerase polyadenylates and releases the nascent mRNA. It then reinitiates RNA synthesis at the next GS signal and caps and methylates the 5' end of the RNA. By responding to the GE and GS signals in this fashion, the polymerase is able to generate subgenomic mRNAs. There is a tendency for the polymerase to dissociate from the template at the gene junctions and, because it can only initiate transcription at the 3' end of the genome, this results in a gradient of expression, with the genes at the 3' end of the genome being transcribed more frequently than those at the 5' end (Collins & Wertz, 1983; Kuo *et al.*, 1996b; Hardy & Wertz, 1998; Krempf *et al.*, 2002; Cheng *et al.*, 2005). In keeping with this gene-expression strategy, the virus genome has evolved such that proteins required in large amounts, such as those involved in disabling innate immune defences, are expressed from the 3' end of the genome, whereas those required in small amounts, such as the polymerase, are expressed from the 5' end (reviewed by Collins *et al.*, 2001).

To replicate the genome, the polymerase binds to a promoter in the Le region and initiates RNA synthesis at the 3' terminus. In this case, as the polymerase moves along the template, it does not respond to the *cis*-acting GE and GS signals and generates a full-length, positive-sense RNA complement of the genome, the antigenome. At the 3' end of the antigenome is the complement of the trailer (TrC), which contains a promoter. The polymerase uses this promoter to generate genome-sense RNA. Unlike mRNA, which is released as naked

RNA, the antigenome and genome RNAs are encapsidated with virus nucleoprotein (N) as they are synthesized and concurrent encapsidation probably accounts for the ability of the polymerase to read through the GE signals during replication.

One of the intriguing questions that this strategy provokes is how the polymerase is controlled between its dual functions of transcription and replication. It is possible that there are two pools of polymerase, i.e. transcriptase and replicase, that have different activities, or that initial interactions of the polymerase with the template commit it to one process or the other. These two possibilities are not mutually exclusive and it may be that both come into play. This question underlies many of the studies on mononegavirus transcription and replication and it is possible that it will only be answered completely when we have a good understanding of the structure of the polymerase and the function of the RNA sequences with which it interacts.

General paradigms of RNA synthesis

In following sections, the *cis*-acting sequences and viral proteins involved in RSV RNA synthesis are described in detail and models for how they might function are discussed. It is useful to give a brief overview of transcription by the well-characterized DNA-dependent RNA polymerases, as these provide an understanding of how polymerase, template and transcript can interact to execute the different events that occur during RNA synthesis and aid interpretation of the data regarding RSV transcription and replication.

Studies on prokaryote transcription have identified the following sequence of events (reviewed by Hsu, 2002; Young *et al.*, 2002; Browning & Busby, 2004; Greive & von Hippel, 2005). First, bonds are formed between specific bases in the promoter and the promoter-binding site on the polymerase-initiation complex, securing the polymerase in place. The polymerase initiates RNA synthesis *de novo* (i.e. without a primer) by positioning the initiating nucleotide and the second NTP in the active site and catalysing formation of the first phosphodiester bond. Polymerization then continues with additional incoming NTPs entering the active site and interacting with the template strand of the DNA. Having initiated RNA synthesis, the polymerase must undergo 'promoter escape', in which it relinquishes the bonds that bind it in place. The transcription complex is highly unstable at this stage and abortive transcripts of ~2–10 nt in length are produced in significant amounts; it is thought that these transcripts are produced by polymerase that remains attached to the promoter and initiates RNA synthesis reiteratively before escaping successfully. Following (or concomitant with) promoter escape, the polymerase changes conformation and enters an elongation mode. Even once the polymerase has begun elongation, there is the possibility that it might stall, terminate (i.e. dissociate), backtrack or continue forward at every NTP-addition cycle. The possibility of each outcome depends on a balance of factors and increasing the rate of transcription and/or stabilizing the elongation complex increases polymerase processivity and promotes elongation. Termination takes place if the complex is destabilized and can occur in response to sequences in the template and/or transcript, and can also be effected by *trans*-acting termination factors.

Termination results in dissociation of the polymerase from the template and release of the nascent RNA.

Although most of our understanding of transcription is derived from studies on prokaryotic polymerases, many of the features described above hold for eukaryotic polymerase II transcription (Dvir, 2002) and experiments on the positive-strand RNA virus *Brome mosaic virus* established that its polymerase follows similar steps in initiation, indicating that these paradigms can be applied to the RNA-dependent RNA polymerases (Sun *et al.*, 1996; Sun & Kao, 1997a, b; Adkins *et al.*, 1998). RNA synthesis by RSV and other members of the *Mononegavirales* is more complex than that described for cellular polymerases, for two main reasons. First, rather than having a promoter for each gene, the virus mRNAs are all produced from a single promoter and their synthesis is dependent on the polymerase stopping and restarting RNA synthesis as it moves along the genome. Second, the genome acts as a template for both transcription and replication and the promoters for these two processes overlap and so, somehow, the polymerase must be governed between initiation of transcription or replication. Therefore, although it is reasonable to assume that RNA synthesis by RSV and other non-segmented, negative-strand RNA viruses can be related to the transcription model described above, their strategy of gene expression and genome replication requires variation on these themes, as discussed below.

The core machinery of RSV transcription and RNA replication

The nucleocapsid template

A key feature of both the genome and antigenome templates is that they remain coated with N protein at all times and so it is the nucleocapsid that is the template for RNA synthesis, rather than naked RNA. This has several benefits for the virus; for example, encapsidation prevents formation of secondary structure in the RNA template, obviating the need for a helicase activity. It also protects the RNA from nuclease attack and reduces production of double-stranded RNA, which might help the virus to avoid provoking antiviral responses (Le Mercier *et al.*, 2002).

The fact that the RNA is coated in N protein raises the question of how the polymerase is able to 'see' and bind to the promoter. Although structural analyses of the nucleocapsid are not yet of sufficient resolution to determine the exact positioning of the RNA relative to N molecules, chemical-modification studies have shown that, in the case of VSV and SeV, the bases of the RNA are exposed in the nucleocapsid, which could potentially allow incoming polymerase complexes to identify the promoter, even though the RNA backbone is encased (Iseni *et al.*, 2000, 2002). In addition, in some paramyxoviruses, the N protein itself might be a component of the promoter. This is suggested by the fact that many paramyxoviruses obligatorily follow the 'rule of six', meaning that their genome nucleotide length must be divisible by 6 to be replicated (Calain & Roux, 1993). This reflects the fact that each molecule of N binds to 6 nt and that the promoter sequence in the RNA is recognized in the context of the molecular signature imparted

by the N protein (Kolakofsky *et al.*, 1998; Vulliémoz & Roux, 2001; Iseni *et al.*, 2002). Thus, in these viruses, the N component of the nucleocapsid might either contribute directly to promoter structure (i.e. by forming a direct interaction with the polymerase) or might position the RNA bases appropriately for recognition by the polymerase. RSV does not follow the 'rule of six' and its genome can be manipulated to be any integer length without loss of template activity (Samal & Collins, 1996). These data indicate that the position of N molecules relative to the sequence in the RNA is not important, suggesting that, although the RSV polymerase recognizes the RNA in the context of the nucleocapsid (and might well interact with the N protein during template binding), there are no specific interactions between polymerase and N protein to position the polymerase over the promoter. Following promoter binding, the RNA template must be brought into the active site of the polymerase. How this happens is not known, but either the structure of the nucleocapsid is such that the polymerase can protrude far enough into it to position the RNA into its active site or the polymerase displaces RNA temporarily from the local N molecules. Electron-microscopy studies have indicated that RSV nucleocapsids are highly flexible structures (Bhella *et al.*, 2002) and this might allow the nucleocapsid to bend to accommodate the large polymerase complex during RNA synthesis.

Although the termini of the RSV genome are complementary, there is no evidence that complementarity is essential for RNA synthesis or that the termini interact, indicating that the RSV template is a linear, rather than panhandle, structure and that the promoter is confined to the 3' end (Peeples & Collins, 2000; Fearn *et al.*, 2002; Cowton & Fearn, 2005). It was found that a critical feature of the nucleocapsid is the structure afforded by its 3' terminus. If the promoter sequence was displaced from the 3' terminus of the nucleocapsid by a small number of nucleotides (≥ 6), mRNA and antigenome synthesis was inhibited almost completely (Collins *et al.*, 1991; Cowton & Fearn, 2005). However, an embedded promoter could be recognized by a polymerase that had already been engaged on the template (Cowton & Fearn, 2005). Thus, it appears that the 3' terminus of the nucleocapsid is important for recruitment of polymerase to the promoter, even though it is not necessary for promoter recognition *per se*. There are at least two possible explanations for this. First, the 3' terminus of the nucleocapsid might combine with the promoter to provide a structure for polymerase assembly. In the case where short (≤ 5 nt) insertions are made between these elements, the promoter complex might expand its footprint to accommodate the 3' terminus and the promoter simultaneously. Second, it is possible that the polymerase interacts first with the structure conferred by the 3' terminus of the nucleocapsid and then displaces the first few N molecules from the end of the RNA in order to interact with the promoter sequence. In this model, proximity of the promoter to the 3' end is important because displacement of N protein can only occur from the terminus, and short 3' extensions are tolerated because each N monomer binds several nucleotides and, so, displacement of one or two N molecules could still reveal the promoter sequence.

The polymerase complex

The minimal complex required for RSV polymerase activity consists of the large protein (L) and phosphoprotein (P) (Mazumder & Barik, 1994). The L–P complex is believed to be responsible for recognition of the promoter, RNA synthesis, capping and methylation of the 5' termini of the mRNAs and polyadenylation of their 3' ends. There is evidence that the L protein contains the enzymic domains required for these processes and its multifunctional nature tallies with its size, with the L gene accounting for almost half of the RSV coding potential (Stec *et al.*, 1991). Alignment of sequences of non-segmented, negative-strand RNA virus L proteins identified six conserved regions (I–VI), which are thought to represent functional domains (Poch *et al.*, 1990). Because of the similarity of these regions between viruses, any function that can be attributed to a conserved region in one virus can probably be extrapolated across the order. On this basis, there is information on which regions of RSV L are involved in polymerization, capping and methylation. Regions II and III probably comprise the polymerization domain, as they contain motifs that are highly conserved in all RNA-dependent RNA polymerases (Poch *et al.*, 1989). The structure of other polymerases has been likened to a right hand, in which the thumb subdomain clamps down onto the RNA template and the palm and finger subdomains form a pocket into which incoming NTPs enter to be added to the growing chain. These subdomains also contain amino acids that are involved in metal-ion coordination, which is necessary for polymerization activity (O'Reilly & Kao, 1998). Although the structure of the polymerase domain has not been determined for any of the negative-sense RNA viruses, their conserved motifs have been modelled onto the known structure of a reverse transcriptase and indicate that the palm and finger subdomains are conserved (Müller *et al.*, 1994).

Evidence that the RSV polymerase might mediate capping was first suggested by the finding that the virus mRNA cap structure is different from that of cellular mRNAs (Barik, 1993). A recent study utilizing an inhibitor that prevents capping of RSV mRNAs attributed guanylyltransferase activity to region V of the L protein and it was proposed that this region acts as a nucleotide-binding domain in guanylylation (Liuzzi *et al.*, 2005). *In silico* analysis predicted that region VI possesses methyltransferase activity (Ferron *et al.*, 2002) and this has now been borne out by functional studies on the VSV and SeV L proteins (Grzelishvili *et al.*, 2005; Li *et al.*, 2005; Ogino *et al.*, 2005).

The phosphoprotein is essential for polymerase activity (Mazumder & Barik, 1994; Barik *et al.*, 1995; Grosfeld *et al.*, 1995; Yu *et al.*, 1995; Dupuy *et al.*, 1999) and, by analogy with other mononegaviruses, might mediate contact between the L protein and the nucleocapsid template (Mellon & Emerson, 1978). A study on fragments of measles virus N and P proteins has shown that the two proteins have fast binding kinetics and weak binding affinity; these features probably allow rapid association and dissociation of the N–P proteins, which could facilitate L–P polymerase movement along the N–RNA template (Kingston *et al.*, 2004). Although there are no similar structural data available for RSV N and P proteins, it is likely that functional features of these proteins are conserved among members of the *Mononegavirales* and that they interact in a similar manner (Karlin *et al.*, 2003).

As its name suggests, the P protein is phosphorylated and there have been several studies to investigate whether there is a role for phosphorylation in coordinating polymerase activity (Barik *et al.*, 1995; Sánchez-Seco *et al.*, 1995). Studies using recombinant viruses or an intracellular plasmid-based assay in which the potential phosphorylation sites on P were abrogated indicated that phosphorylation is not essential for P protein function (Villanueva *et al.*, 2000; Lu *et al.*, 2002). However, a study using an *in vitro* RSV RNA-synthesis assay provided a different result. In this study, if phosphorylation at the major acceptor site (serine 232) was abolished, the polymerase was still able to initiate RNA synthesis at the 3' terminus of Le, but only produced short transcripts, predominantly 9 or 11 nt long (Dupuy *et al.*, 1999). Given that the RSV promoter is likely to consist of Le nt 1–11 (see below), these are the correct length to be abortive transcripts, indicating that P phosphorylation might be necessary to allow the polymerase to navigate promoter escape. It is currently unclear why the intracellular and *in vitro* studies give apparently different results. One possibility is that cellular factors can complement phosphorylated P activity. Another possibility is that the promoter sequences used were slightly different, which could potentially affect the efficiency of promoter escape.

Although the L–P complex is sufficient to perform RNA synthesis *in vitro* and can be considered as the core polymerase, this complex is unlikely to function independently in nature. As described above, cellular proteins actin and profilin are necessary for optimal RNA-synthesis activity and could potentially be part of the polymerase complex (Burke *et al.*, 1998, 2000). The viral M2-1 protein is required for processivity during transcription and is likely to be a component of the 'transcriptase' complex (Collins *et al.*, 1996). The viral N protein is required to encapsidate the nascent RNA during replication and it is possible that the 'replicase' consists of a complex of L–P–N, as has been shown for VSV (Qanungo *et al.*, 2004).

Transcription of mRNAs

Models for transcription initiation in the *Mononegavirales*

The Le sequence at the 3' end of the genome is highly complex, as it is important for both transcription and replication initiation. How it is able to direct two processes is not known and determining the site of transcription initiation is key to understanding this. It is known that almost all mRNA from the first gene of the RSV genome is initiated at the first GS signal at position 45 (Collins & Wertz, 1985). However, the mechanism by which the polymerase initiates at this site is poorly understood and remains controversial. There are two basic models for transcription initiation in the order *Mononegavirales*, based mainly on studies with SeV and VSV. Model 1 proposes that the polymerase first initiates RNA synthesis opposite the first nucleotide of the genome and transcribes the Le region. The polymerase then releases this transcript at, or near, the end of the Le region, locates the first GS signal and uses this to reinitiate RNA synthesis. Model 2 proposes that the polymerase initiates transcription directly at the GS signal. According to this model, there are two pools of polymerase, dedicated to either replication or transcription, that recognize distinct initiation sites. There are a number of studies

that support model 1 (reviewed by Kolakofsky *et al.*, 2004); for example, reverse-genetics analyses with SeV indicate that the nature of the promoter sequence can affect how efficiently the polymerase can initiate at the GS signal, which indicates that the two events are linked (Vulliémoz & Roux, 2002; Le Mercier *et al.*, 2003; Vulliémoz *et al.*, 2005). Model 2 is supported by studies with VSV that indicate that initiation at the first GS signal occurs independently of Le transcription (Chuang & Perrault, 1997; Whelan & Wertz, 2002). Arguments have been made against these conclusions (Kolakofsky *et al.*, 2004), but compelling evidence for model 2 is derived from a study of VSV RNA synthesis *in vitro* using purified polymerase, which showed that two pools of polymerase complex could be identified that were associated with different proteins. In this study, the replicase and transcriptase complexes were found to initiate at the 3' end of Le or directly at the GS signal, respectively (Qanungo *et al.*, 2004). Presumably, in this case, the transcriptase either binds directly at the GS signal or scans from the 3' end of the nucleocapsid to the GS signal without initiating RNA synthesis. Currently, it is difficult to reconcile the data supporting the two models. One possibility is that paramyxo- and rhabdoviruses use different initiation mechanisms and, in this respect, it is noteworthy that the organization of the RSV and VSV promoter regions appears to be different (Li & Pattnaik, 1999; Whelan & Wertz, 1999; McGivern *et al.*, 2005). Alternatively, it is possible that both transcription-initiation mechanisms come into play during infection and so experimental evidence can be obtained to support both models. The *cis*-acting signals for RSV transcription initiation have been broadly mapped and, although they do not distinguish between these two models, they do provide information on the initial interactions of the polymerase with the template.

Initiation of transcription at the 3' end of the RSV genome

RSV transcription initiation involves three sequence elements: a region at the 3' end of the Le, a U-rich region at the end of the Le and the first GS sequence (indicated in bold type in Fig. 2a). As yet, there are no biochemical data regarding the functions of these sequences in RSV, but reverse-genetics analysis has suggested what roles they could play.

Saturation-mutation analysis of the first half of the RSV Le identified a number of nucleotides near the Le 3' terminus as being important for mRNA synthesis: substitutions at positions 3, 4, 5, 8, 9, 10 and 11 reduced mRNA production to <30 % of wild-type levels (Fearn's *et al.*, 2002; McGivern *et al.*, 2005). Thus, regardless of the site of transcription initiation, mRNA synthesis from the GS signal is dependent on a sequence at the 3' end of the Le. There is evidence that this sequence is a polymerase-binding site. As described above, it was shown that the Le sequence must be at, or near, the 3' end of the template for RNA synthesis to occur (Collins *et al.*, 1991; Cowton & Fearn's, 2005). Further analysis to determine how much of Le needed to be proximal to the 3' terminus showed that only the first 11 nt of the Le were required. Therefore, this finding suggests that nt 1–11 contain a polymerase-binding site that functions cooperatively with the 3' terminus of the nucleocapsid to recruit transcriptase to the template (Cowton & Fearn's, 2005).

The first GS signal (the NS1 GS sequence) is also integral to the sequential-transcription programme (Fig. 2a). If this sequence is ablated, a small amount of transcript corresponding to the first gene is produced, but it is initiated at the 3' terminus of the Le and downstream genes are transcribed at only approximately 10 % of wild-type levels (Kuo *et al.*, 1996a). Thus, the GS signal is required for accurate and efficient transcription initiation. The sequence of the 3'-proximal GS signal is the same as most of the other RSV GS signals and this sequence responded similarly to saturation mutagenesis, irrespective of whether it was at the start of the first or second gene (Kuo *et al.*, 1997). These data indicate that the 3'-proximal GS sequence is functionally equivalent to the internal GS signals, discussed in detail below.

Although the 3'-terminal 11 nt and GS sequence are sufficient to signal transcription initiation, alone they are inefficient. Efficient transcription depends on a region at the end of Le, immediately upstream of the GS signal (Fig. 2a; Fearn's *et al.*, 2000; McGivern *et al.*, 2005). Although this has not been mapped precisely, it contains a stretch of uridylates that are conserved in the Le regions of all pneumoviruses sequenced to date (Fig. 2b), suggesting that the U-rich nature of this sequence is important. The function of this sequence is not known, but in VSV, it has been shown that GS-signal activity is dependent on the presence of an upstream U tract (Hinzman *et al.*, 2002).

If the U-rich region and GS signal are moved closer to the element at the 3' end of Le, then transcription is inhibited significantly. The signals can still function if they are moved further apart; for example, a 60 nt insertion upstream of the U-rich region is tolerated, but insertions of 80 nt or more are inhibitory (Fearn's *et al.*, 2000; McGivern *et al.*, 2005). These data suggest that there is some latitude in the spacing between the 3' element and the U-rich region and GS signal; however, they cannot function if they are brought closer together and they cannot be displaced too far.

These findings do not distinguish between the two models of transcription, although they do indicate that the transcriptase is recruited to the 3' terminus of the genome, rather than directly at the GS site. The polymerase presumably then moves to the GS signal to initiate mRNA synthesis. The possible mechanisms by which the polymerase could migrate from the 3' terminus to the GS signal are discussed in detail in a later section. Probably the signals cannot function well if they are closer together because polymerase pausing at the GS signal inhibits polymerase binding to the signal at the 3' end.

Initiation of mRNA synthesis at the GS signals

The GS signals direct the polymerase to initiate mRNA synthesis. These are highly conserved sequences, with only one, the L GS signal, differing from the sequence 'CCCCGUUUUAU' (Collins *et al.*, 1986, 1987). It is not clear exactly how the GS signals are able to direct RNA-synthesis initiation, as they do not act as promoters to recruit exogenous polymerase to the template. However, it is possible that the sequence can act similarly to a promoter for a polymerase that is already in a template-bound conformation, forming specific contacts that hold the polymerase in place for sufficient time to allow the initiating nucleotide to enter the polymerase active site and act as a substrate for RNA-synthesis initiation.

As described above, there is evidence that the virus L protein is responsible for capping the mRNA transcripts (Liuzzi *et al.*, 2005). Uncapped transcripts are truncated after approximately 50 nt, suggesting that addition of the cap structure allows the polymerase to change into a stable elongation mode (Liuzzi *et al.*, 2005). This is reminiscent of RNA polymerase II transcription, in which cap addition and elongation are linked, providing a checkpoint mechanism to ensure that only capped transcripts are extended (reviewed by Orphanides & Reinberg, 2002; Shilatifard, 2004; Zorio & Bentley, 2004). Studies with VSV and SeV indicate that both the capping and methylation reactions are dependent on interplay between the sequence encoded by the GS signal and the virus polymerase (Rose, 1975; Stillman & Whitt, 1999; Ogino *et al.*, 2005). Saturation-mutagenesis analysis of the RSV GS sequence showed that it is highly stringent, with most substitutions causing a significant decrease in gene expression (Kuo *et al.*, 1997). This rigidity is striking and may reflect the possibility that the RSV GS sequence is important for modifying the 5' end of the nascent RNA, in addition to initiation of mRNA synthesis.

Sequence analysis of transcripts from a mutant RSV GS signal revealed a surprising phenomenon termed 'quasi-templating'. A GS signal containing a substitution at position 1 generated a mixed population of mRNA transcripts: two-thirds contained the exact complement of the mutant GS sequence, but one-third contained a wild-type 'G' residue at the first position (Kuo *et al.*, 1997). Thus, it appears that the 'correct' nucleotide can be inserted at the first position of the transcript, even if this is contrary to what is being directed by the template. It should be noted that, although one-third of transcripts showed evidence of quasi-templating in this study, this might not represent one-third of initiations. The position 1 GS mutants only produced low levels of detectable full-length mRNA, indicating that the mutation hindered some aspect of mRNA synthesis. If the nucleotide at position 1 was important for capping of the mRNA, any transcripts containing the 'correct' non-templated nucleotide at this position would be more likely to be extended to the end of the gene than transcripts containing the mutant assignment. Thus, in this scenario, transcripts in which the first position was quasi-templated would be enriched, suggesting a deceptively high level of quasi-templating. Nonetheless, even if this caveat is taken into account, the evidence suggests that at least 3 % of initiations are quasi-templated, which is significantly higher than the expected error rate of an elongating polymerase ($\leq 10^{-3}$). This raises the possibilities that the polymerase is either very error-prone during initiation of mRNA synthesis, that it preferentially uses a G residue as the initiating nucleotide or that it can initiate by using a prime–realign mechanism.

Elongation of the mRNA transcripts: the role of M2-1 in transcriptase processivity

Having initiated RNA synthesis, capped the mRNA and entered an elongation mode, the transcriptase remains subject to destabilization. It can terminate prematurely at multiple intragenic sites, presumably in response to sequences in the template and/or transcript that undermine the integrity of the elongation complex (Grosfeld *et al.*, 1995; Collins *et al.*, 1996; Fearn & Collins, 1999b). This is prevented from occurring by the M2-1 processivity factor. Although the main function of M2-1 is to prevent inappropriate intragenic termination, it can

cause the polymerase to read through the GE sequences and this is presumably another manifestation of the same processivity function (Hardy & Wertz, 1998; Hardy *et al.*, 1999; Sutherland *et al.*, 2001). A direct comparison of the intragenic and GE anti-termination activities indicated that a higher concentration of M2-1 was required for processivity through a GE sequence, which is consistent with this view (Fearn & Collins, 1999b). Because transcription of RSV genes is dependent on complete and sequential transcription of each upstream gene, M2-1 activity is necessary not only to allow the polymerase to transcribe to the end of each gene, but also to access the promoter-distal regions of the genome. In addition, its ability to cause readthrough at the intergenic regions could alleviate the transcription gradient, as discussed previously by Hardy & Wertz (1998) and Whelan *et al.* (2004). The M2-1 protein has a Cys₃-His₁ motif near its N terminus that has been proposed to interact with zinc, allowing appropriate folding of the M2-1 protein and binding to RNA (Hardy & Wertz, 2000; Cartee & Wertz, 2001; Tang *et al.*, 2001; García-Barreno *et al.*, 2005). In addition, M2-1 binds to P and this is necessary for M2-1 function (Mason *et al.*, 2003). Early studies suggested an interaction between N and M2-1, but this has since been shown to be an artefact of their RNA-binding activities (Cartee & Wertz, 2001). Thus, it seems likely that M2-1 is a component of the transcriptase complex and interacts with the P moiety of the polymerase and either the nascent or template RNA. There are data indicating that M2-1 binds specifically to RNA sequence encoded by the Le region, suggesting that M2-1 might be 'loaded' onto the polymerase during transcription initiation (Cuesta *et al.*, 2000). However, M2-1 activity is evident with minigenomes that have large substitutions in the Le region (Collins *et al.*, 1991; McGivern *et al.*, 2005). Therefore, this interaction either reflects specificity for the first 11 nt encoded by Le, which were conserved in these mutants, or a preference for binding to AU-rich RNA.

Transcription termination at the GE signals

At the end of each gene is a GE signal consisting of a conserved sequence followed by an uridylyate tract, which signals the polymerase to polyadenylate and release the nascent mRNA. It has not been determined how the RSV GE signal controls these events; however, extensive functional analysis of the VSV GE signal has been performed and these findings provide a model for RSV termination. The VSV GE sequence is 'AUACUUUUUUU' and the AUAC sequence functions together with the U tract to cause the polymerase-transcript complex to undergo repetitive cycles of slippage relative to the template, whilst the polymerase active site is positioned on the U tract, resulting in polyadenylation of the mRNA transcript. It is currently unclear how release of the mRNA after a certain number of recurring adenylation cycles is controlled, but there is evidence to suggest that the poly(A) tail of the nascent mRNA is a factor in the release process, which might provide a mechanism to control poly(A) tail length (reviewed by Barr *et al.*, 2002). In VSV, the GE signal is not the only important factor in termination, as there is evidence of an intimate link between 5' modification of the mRNA and correct polyadenylation and release, indicating that termination is a complex process (Rose *et al.*, 1977; Whelan *et al.*, 2000).

In contrast to the VSV GE signals, those of RSV are somewhat divergent and vary in their termination efficiencies (Kuo *et al.*, 1997; Hardy & Wertz, 1999; Tran *et al.*, 2004). Most are not maximally efficient and, so, readthrough mRNAs representing two or more genes are found in infected cells (Collins & Wertz, 1983). The GE signal of RSV (A2 strain) consists of a UCAAU motif that is conserved in nine of the ten GE signals, a variable region of 3 or 4 nt, mostly consisting of adenylate and uridylate residues, and a tract of at least four uridylate residues. Each of these elements is important for GE function and the nucleotides immediately adjacent to the GE signal can also affect termination efficiency (Harmon *et al.*, 2001; Sutherland *et al.*, 2001; Harmon & Wertz, 2002; Cartee *et al.*, 2003; Moudy *et al.*, 2003). It is likely that the U tract of the GE signal is important for polyadenylation and that the A/U nature of the sequence contributes to polymerase instability (Whelan *et al.*, 2004).

Two RSV L mutants that arose serendipitously have suggested a specific role for the polymerase in transcription termination. Each mutant contains a single amino acid substitution, one at aa 1049 and the other at aa 1169. The mutant polymerases appear to be functional in every respect, except that they demonstrate high levels of readthrough at the gene junctions and thus are somewhat defective in transcription termination (Juhász *et al.*, 1999; Cartee *et al.*, 2003). This raises a question as to what these mutations are altering. One possibility is that the wild-type polymerase has evolved such that its processivity during transcription is not maximal, but is finely tuned to ensure adequate termination at the GE signals, and that the mutations alter polymerase structure such that its processivity is increased. However, the data for at least one of the mutants suggest that this is not the case. The 1049 mutant polymerase displays a dependence on M2-1 to synthesize complete mRNA transcripts similar to that of the wild-type polymerase, demonstrating that it is not more processive (Cartee *et al.*, 2003). Thus, it appears that the mutation affects termination specifically at the GE signals and not termination *per se*, indicating that termination at the GE signals involves a specific site on the polymerase and is not simply due to instability between the polymerase active site, template and transcript.

Evidence for polymerase scanning from GE to GS

In other polymerases, transcription termination results in release of the polymerase complex from the transcript and template. However, in the order *Mononegavirales*, most of the polymerase that terminates at the GE signal releases the transcript without dissociating from the template. Because the polymerase has released the mRNA, its active site is free to reinitiate mRNA synthesis if it encounters a GS signal. At most RSV gene junctions, there is an intergenic region between the GE and GS signals and the polymerase is thought to scan this to locate the next GS signal.

The RSV intergenic regions vary from 1 to 56 nt in length and diverge between different virus strains (Collins *et al.*, 1986; Johnson & Collins, 1988; Tolley *et al.*, 1996). Other than the nucleotides immediately flanking the GE and GS sequences, there are no obvious sequence motifs in the intergenic regions except that they are highly U-rich, particularly at the end of the longer intergenic regions. Different intergenic regions could be exchanged for each other with no significant effect on transcription of the flanking genes (Kuo *et al.*, 1996b) and variations in

the intergenic regions of clinical isolates did not have an obvious effect on transcription of the downstream gene (Moudy *et al.*, 2004). These data suggest that the intergenic regions do not fulfil an essential function in gene expression, although substitution of a wild-type intergenic region with an artificial G/C-rich sequence caused a 30 % reduction in expression of the downstream gene, demonstrating that certain sequences can be inhibitory to reinitiation (Kuo *et al.*, 1996b).

The fact that the intergenic regions are variable in length indicates that polymerase can scan the template from GE to GS sequences and there is evidence to suggest that this scanning activity is highly efficient. For example, a downstream gene was still expressed even if the intergenic region was extended to 612 nt, although longer intergenic regions could not be tolerated (Bukreyev *et al.*, 2000; A. Bermingham & P. Collins, personal communication). A study of the unusual M2/L gene junction, in which the GS signal for the (downstream) L gene lies upstream of the M2 GE signal, also provides evidence of polymerase scanning (Collins *et al.*, 1987; Fearn & Collins, 1999a). The arrangement of this junction is shown in Fig. 3. Reverse-genetics analysis indicated that the polymerase transcribes M2 to the GE signal and then scans backwards to reinitiate at the L GS site. It was also shown that, having reached a GE signal, the polymerase could reinitiate RNA synthesis from either an upstream or a downstream GS signal. Thus, it appears that, having released mRNA at a GE signal, the polymerase can maintain a stable interaction with the template and scan in either direction to locate a GS signal.

A number of cellular proteins are capable of scanning for a specific sequence within a nucleic acid backbone and these provide a model for how RSV polymerase scanning might occur. One of the best characterized is the *Escherichia coli lac* repressor protein. The *lac* repressor operator DNA-target site contains a sequence that allows formation of specific hydrogen bonds between the repressor and the DNA. To locate this sequence, the *lac* repressor diffuses along DNA by virtue of electrostatic interactions with the sugar–phosphate backbone. When it encounters its binding site, the protein undergoes a conformational change and hydrogen bonds are formed with the bases in the DNA, holding the repressor in place (von Hippel, 2004). In the case of RSV, a scanning model must take the encapsidated nature of the template into consideration. As described above, the nature of the P–N interactions may allow rapid movement of the polymerase along the nucleocapsid and, as the whole RNA genome is coated with N protein, it would be expected that there is no net loss or gain of energy as P–N interactions are broken and reformed (Kolakofsky *et al.*, 2004). Thus, once the polymerase has released the mRNA of an upstream gene, it could diffuse along the template until it reaches the GS signal and forms specific contacts with the bases in this sequence. If this model is correct, the artificial, G/C-rich intergenic sequence that inhibited reinitiation at a downstream GS signal might have introduced hydrogen-bond donors or acceptors that mimicked those in the GS sequence and stalled the polymerase before it reached the authentic GS signal.

Having reinitiated mRNA synthesis, the polymerase can continue along the template, elongating to the end of a gene, polyadenylating and releasing the mRNA and then scanning and reinitiating mRNA synthesis at the start of the next gene, in sequence until it reaches the end of the template. As indicated above, this cycle of RNA synthesis is not completely efficient

and the polymerase can dissociate from the template, resulting in a gradient of gene expression. In VSV, polymerase dissociation was found to occur mainly at, or near, the gene junctions (Iverson & Rose, 1981) and presumably occurs either due to dissociation during termination at the GE signals, failure to locate and initiate at the GS signal, failure to cap the mRNA or a combination of the above.

In summary, the data suggest that, during RSV transcription, the polymerase is recruited to the 3' end of the nucleocapsid, translocates to the first GS signal and initiates mRNA synthesis. From this point, RSV transcription is similar to transcription by cellular polymerases, with GS signals acting like promoters to template-bound polymerase, and polymerase processivity and termination being dependent on a combination of *cis*- and *trans*-acting factors. A difference between RSV and cellular transcription lies in termination, with a large proportion of the RSV polymerase maintaining a template-bound conformation and entering a scanning mode following mRNA release.

Genome replication

One fundamental difference between transcription and replication is that, during replication, the polymerase fails to respond to the GE signals, allowing it to synthesize a complete complement of the genome. The second significant difference is that the replication product is encapsidated as it is synthesized and it is likely that these two features are linked. Therefore, in considering replication initiation, it is important to consider not only how RNA synthesis is initiated, but also how the nascent RNA might become encapsidated.

***cis*-acting signals for antigenome initiation**

The first 34 nt of the Le region contain all the sequences required for antigenome synthesis (indicated in bold type in Fig. 4a), setting RSV and other pneumoviruses apart from other members of the family *Paramyxoviridae*, which have bipartite replication promoters that extend into the first gene (Collins *et al.*, 1991; Murphy *et al.*, 1998; Tapparel *et al.*, 1998; Murphy & Parks, 1999; McGivern *et al.*, 2005). Saturation-mutagenesis analysis of the Le region showed that nt 3, 4, 5, 8, 9, 10 and 11, which are important for transcription, are also necessary for replication and, with the exception of one substitution at position 4, substitutions at these positions affected mRNA and antigenome synthesis similarly, suggesting that these nucleotides play a common role in both processes. In addition, mutations at positions 1, 2, 6 and 7 inhibited antigenome production significantly, although they had a relatively minor effect on mRNA synthesis, indicating that these nucleotides play a role specifically in replication (Fearn *et al.*, 2002). By using primer-extension analysis to detect RNA transcripts initiated at the Le 3' terminus, it was determined that the first 13 nt of Le are sufficient to direct initiation of RNA synthesis (Cowton & Fearn, 2005; McGivern *et al.*, 2005) and, similarly to the findings for transcription, there is evidence that replicase is recruited by the first 11 nt of Le in conjunction with the 3' end of the nucleocapsid. Together, these data indicate that Le nt 1–11 contain a

binding site for the replicase and can function as a promoter to signal antigenome synthesis. Currently, it is not known whether nt 3, 4, 5, 8, 9, 10 and 11 are sufficient to signal replication initiation or whether nt 1, 2, 6 and 7 are also required.

Although the 3' terminus of the nucleocapsid is clearly important for polymerase recruitment to the template, it apparently does not play a major role in determining the antigenome start site. When antigenome transcripts from minigenomes containing short 3'-terminal extensions were sequenced, it was found that they were initiated predominantly opposite the first nucleotide of the Le sequence, rather than opposite the first nucleotide of the template (Cowton & Fearn, 2005). Thus, it appears that, although the promoter sequence must be close to the 3' end of the nucleocapsid to recruit polymerase to the template, the promoter sequence alone is sufficient to direct initiation opposite the first nucleotide of Le, indicating that interaction with this sequence positions the polymerase active site appropriately for antigenome initiation.

Encapsidation of antigenome RNA

Studies with VSV and SeV showed that the virus N protein is required *in trans* for virus RNA replication, probably due to its role in encapsidation (Patton *et al.*, 1984; Horikami *et al.*, 1992). Concurrent encapsidation increases polymerase processivity during SeV replication and it is likely that increased processivity allows the polymerase to override the GE termination signals to generate a full-length antigenome (Vidal & Kolakofsky, 1989; Gubbay *et al.*, 2001). The mechanism by which RSV encapsidation occurs is not completely understood, but parts of the jigsaw are now beginning to come together.

If expressed in isolation, the RSV N protein binds non-specific RNA to form nucleocapsid-like structures, demonstrating that N–N interactions are sufficient to drive nucleocapsid assembly (Bhella *et al.*, 2002; Murphy *et al.*, 2003). In infected cells, most soluble N protein is found in a complex with P, which is thought to prevent N binding to non-specific RNA (Cuesta *et al.*, 2000; Murphy *et al.*, 2003; Castagné *et al.*, 2004). Consistent with this hypothesis, there is evidence that P interaction with N interferes with N–RNA binding directly (García-Barreno *et al.*, 1996; Khattar *et al.*, 2000; Murphy *et al.*, 2003). Thus, presumably encapsidation involves a conformational change in the N–P complex that allows N to be released from P, exposing its RNA-binding site.

There is evidence that a *cis*-acting signal directs initiation of encapsidation. Initiation products generated from minigenomes with mutations in the central part of Le were not encapsidated as efficiently as those generated from minigenomes containing wild-type sequence. In addition, although nucleotides at the 3' end of Le were found to be sufficient to direct efficient initiation of antigenome synthesis, production of full-length antigenome also required the central region of Le (Fig. 4; McGivern *et al.*, 2005). These data suggest that the central part of Le contains sequences that facilitate antigenome encapsidation, which in turn enables the polymerase to be processive through the gene junctions to the end of the template. The sequence required for encapsidation and processivity has not been mapped precisely and

could include nucleotides at the 3' end of Le, including nt 1, 2, 6 and 7, which are more important for replication than transcription.

This encapsidation signal could function in various ways. One possibility is that the sequence encodes a signal at the 5' end of the antigenome RNA, which recruits N–P complexes and causes N to be delivered to the nascent RNA chain. This could then act as a nucleation site to which further N molecules are delivered. Alternatively, the encapsidation signal could function in the context of the template; for example, it might recruit factors that facilitate encapsidation or cause the polymerase to adopt a conformation that allows it to bind N (or N–P), which results in delivery of N to the nascent RNA as it is being extruded.

Initiation of genome RNA synthesis

The TrC promoter region at the 3' end of the antigenome is 155 nt in length (Mink *et al.*, 1991). This promoter directs a higher level of replication than the Le, resulting in an imbalance between antigenome and genome in infected cells. By analogy with *Rabies virus*, this bias might be important to promote incorporation of nucleocapsids of genome polarity into particles during virus assembly (Finke & Conzelmann, 1997). The TrC promoter region has not been mapped as well as the Le, but some information is available, due to the similarities between the two sequences. As shown in Fig. 4(b), the RSV TrC and Le are highly conserved for the first 26 nt, but differ thereafter. As the high level of conservation would suggest, the 3' ends of the Le and TrC regions appear to be functionally similar: a direct comparison of the first 34 nt of Le with the first 36 nt of TrC indicated that these promoter regions have similar strengths and activities, although the TrC sequence directed a slightly higher level of replication, indicating that it is optimized for this process, whereas the Le is not (Fearn's *et al.*, 2000). Single nucleotide substitutions in positions 1–7 of TrC behaved similarly to the corresponding substitutions in Le, suggesting that the remainder of the Le and TrC sequence does not alter how this 3'-terminal sequence functions (Peeples & Collins, 2000). However, there is evidence that positions 4 and 12 are optimized for transcription in the Le promoter and for replication in the TrC promoter (Fearn's *et al.*, 2002). These differences in replication efficiency could reflect differences in replicase binding, initiation or encapsidation.

The reason why the TrC promoter region directs a significantly higher level of replication than the Le is primarily due to sequence lying between TrC nt 36 and 155 (Fearn's *et al.*, 2000). This observation was originally made by using a minigenome and has been confirmed in a more authentic setting by introducing deletions into the 5' Tr region of an infectious clone of RSV such that the antigenomic promoter was shortened to the 3' 36 or 57 nt. The mutant viruses were able to replicate, but generated lower levels of genome RNA than wild-type virus, confirming that the first 36 nt of the TrC promoter region are sufficient to direct replication initiation and encapsidation, but that the remainder of the TrC region enhances replication (R. Fearn's & P. Collins, unpublished data). The mechanism by which this enhancing sequence functions is currently not known.

How do the data for RSV RNA synthesis integrate into the models for transcription and replication initiation?

As described above, there are two models for replication and transcription initiation, one in which both processes initiate at the 3' end of the Le (model 1) and one in which the transcriptase and replicase initiate at different sites (model 2). Although the data available for RSV do not confirm either model, they do provide information that allows these models to be fine-tuned.

Based on the data described, the first 11 nt of Le appear to contain both the replicase- and transcriptase-binding sites. There is evidence that these nucleotides also act as the replication promoter and are sufficient to direct initiation opposite nt 1. This suggests that, when the polymerase is recruited to the template, its active site is positioned over nt 1. Thus, to initiate mRNA synthesis at the GS site, the polymerase must move relative to the template in order to position its active site over the GS sequence at nt 45. Evidence that the transcriptase can do this was derived from experiments suggesting that polymerase recruited to nt 1–11 at the 3' end of the template can scan to access internal sequences (Cowton & Fearn, 2005). If this scenario is correct, then the polymerase must break its bonds with its binding site at the 3' end of the nucleocapsid and it would be expected that this would require energy.

The contacts between a polymerase and its binding site are typically broken once the polymerase has initiated RNA synthesis. Thus, model 1, in which transcription is initiated at position 1 of the Le, provides a simple explanation of how the polymerase accesses the GS signal (Fig. 5a). As discussed previously by Kolakofsky *et al.* (2004), release of the Le transcript near the end of Le could occur as a consequence of failure to modify the 5' terminus of the RNA. If capping was dependent on the sequence encoded by the GS signal, the RSV Le transcript would lack the sequence motif required, causing the polymerase to release the transcript after approximately 50 nt. The polymerase active site would then be unoccupied, allowing reinitiation at the GS signal. For replication to occur, the nascent RNA must become encapsidated before the polymerase reaches the 50 nt threshold. The data discussed here indicate that this could occur, as all of the sequences required for encapsidation of the RNA are located within the first 34 nt of the Le region. If encapsidation was initiated, then the processivity of the polymerase would be increased, allowing it to continue beyond 50 nt, read through the GE signals and produce an antigenome (Fig. 5c). The data regarding the role of P phosphorylation in promoter escape are consistent with this model: in transcription reactions in which a non-phosphorylated P protein was used, short transcripts of approximately 11 nt were synthesized, but there was no evidence for initiation of mRNA synthesis at the GS signal (Dupuy *et al.*, 1999). These data suggest that the polymerase initiates and elongates transcripts from the promoter at the 3' end of Le before initiating mRNA synthesis at the first GS signal.

If RSV transcription is initiated directly at the GS signal, as in model 2, then the polymerase needs to break contact with nt 1–11 without initiating RNA synthesis. It is possible that it is able to hydrolyse ATP to do this (Fig. 5b). Following release of the polymerase from its binding site, it would be able to scan along the template by using the same mechanism as when scanning at the gene junctions. When it reached the GS signal, it would be directed to initiate

RNA synthesis. According to this model, because replication is initiated at position 1 of the Le, it would not require an ATPase activity. In the case of VSV, it has been shown that a higher concentration of ATP is required for transcription than for replication initiation, which is consistent with this model (Testa & Banerjee, 1979; Perrault & McLear, 1984; Beckes *et al.*, 1987). Similarly to model 1, because the signals for encapsidation are within the Le region, the replication product could become encapsidated, whereas the transcription product would not be and any replication products that failed to become encapsidated would be aborted, due to poor processivity. In both of these models, the U-rich region at the end of Le might stall the polymerase to allow it to recognize the GS signal efficiently.

Regulation of transcription and replication

The processivity factors: M2-1 and N-P

As described above, M2-1 and N-P increase polymerase processivity and can cause the polymerase to read through the GE signals. Thus, these proteins were likely candidates to regulate the polymerase between its transcription and replication activities. However, this does not appear to be the case. Increasing the level of intracellular M2-1 has no detectable effect on antigenome or genome synthesis or on the levels of transcription initiation (Collins *et al.*, 1996; Hardy & Wertz, 1998; Fearn & Collins, 1999b), demonstrating that this protein is not involved in modulating the polymerase between its dual activities. Increasing the intracellular concentration of N and P results in an increase in replication, but does not inhibit transcription (Fearn *et al.*, 1997). At face value, this suggests that mRNA and antigenome syntheses occur independently, consistent with model 2 described above. However, if model 1 is correct, it is possible that only a fraction of initiations at the 3' end of the Le result in successful reinitiation at the GS signal or encapsidation. Thus, an increase in encapsidation, and the consequent increase in antigenome synthesis, might not have any detectable impact on transcription.

The M2-2 protein

One protein that might regulate RSV polymerase activity is the virus M2-2 protein. M2-2 has a potent negative effect on RSV RNA synthesis and its expression is moderated in a natural virus infection (Collins *et al.*, 1996; Hardy & Wertz, 1998; Bermingham & Collins, 1999; Ahmadian *et al.*, 2000; Cheng *et al.*, 2005). If M2-2 expression is ablated in recombinant RSV, the relative levels of RSV replicative and mRNAs are altered (Bermingham & Collins, 1999; Jin *et al.*, 2000), suggesting that M2-2 promotes antigenome and genome accumulation, but inhibits mRNA synthesis. It should be noted that, although M2-2 might modulate the polymerase between its different activities, it cannot be the only factor that does this, as the mutant virus produces both replication and transcription products. The mutant virus mediated enhanced cell-cell fusion compared with its wild-type counterpart, which could reflect a defect in virus packaging and release. In addition, M2-2 was found to augment release of infectious virions in a

reconstituted packaging assay (Teng & Collins, 1998). These data suggest that the M2-2 protein could regulate nucleocapsid RNA synthesis in preparation for virus assembly.

Concluding remarks

In this review, we have described the elegant and economical strategy that RSV uses to transcribe and replicate its genome. Transcription and replication are initiated by using overlapping, *cis*-acting sequences, the transcription strategy represents a simple means of generating multiple mRNA without the necessity for numerous promoters and replication is modulated such that synthesis of genome-sense RNA predominates. As we have attempted to convey, the paradigms that have been established with other RNA polymerases provide a useful platform for interpreting the data concerning RSV RNA synthesis and give hints as to the molecular details that are involved. However, in contrast to other organisms, RNA synthesis by RSV and other members of the order *Mononegavirales* has evolved to entail numerous signalling interactions between the polymerase complex, associated proteins, template and transcript, which allow the production of different RNAs.

Our understanding of the way in which the RSV genome is expressed and replicated is already leading to potential benefits in controlling the virus. Knowledge of the roles of the Le, Tr and gene-junction sequences and the identification of the polymerase and its associated proteins led to the creation of the infectious clone of RSV, which has allowed rational design of live-attenuated vaccine candidates (Collins *et al.*, 1999; Collins & Murphy, 2002). The unique nature of the RSV polymerase makes it an ideal target for antiviral therapy and inhibitors are now being identified that can block its activity (Mason *et al.*, 2004; Liuzzi *et al.*, 2005). We are now in a position to develop an understanding of the molecular interactions between the components involved in RNA synthesis that, in the future, may open more avenues for controlling RSV and related viruses.

Acknowledgements

We thank Peter Collins and Alison Bermingham for providing unpublished information, Dave Bhella, Kirsty MacLellan, Richard Sugrue and Sean Whelan for helpful discussion and Julie Bracken for critical reading of the manuscript. The work in the authors' laboratory is funded by The Wellcome Trust, the Biotechnology and Biological Sciences Research Council and the Medical Research Council (UK).

References

- Adkins, S., Sawicki, S. S., Faurote, G., Siegel, R. W. & Kao, C. C. (1998).** Mechanistic analysis of RNA synthesis by RNA-dependent RNA polymerase from two promoters reveals similarities to DNA-dependent RNA polymerase. *RNA* **4**, 455–470.
- Ahmadian, G., Randhawa, J. S. & Easton, A. J. (2000).** Expression of the ORF-2 protein of the human respiratory syncytial virus M2 gene is initiated by a ribosomal termination-dependent reinitiation mechanism. *EMBO J* **19**, 2681–2689.
- Barik, S. (1992).** Transcription of human respiratory syncytial virus genome RNA in vitro: requirement of cellular factor(s). *J Virol* **66**, 6813–6818.
- Barik, S. (1993).** The structure of the 5' terminal cap of the respiratory syncytial virus mRNA. *J Gen Virol* **74**, 485–490.
- Barik, S., McLean, T. & Dupuy, L. C. (1995).** Phosphorylation of Ser²³² directly regulates the transcriptional activity of the P protein of human respiratory syncytial virus: phosphorylation of Ser²³⁷ may play an accessory role. *Virology* **213**, 405–412.
- Barr, J. N., Whelan, S. P. J. & Wertz, G. W. (2002).** Transcriptional control of the RNA-dependent RNA polymerase of vesicular stomatitis virus. *Biochim Biophys Acta* **1577**, 337–353.
- Beckes, J. D., Haller, A. A. & Perrault, J. (1987).** Differential effect of ATP concentration on synthesis of vesicular stomatitis virus leader RNAs and mRNAs. *J Virol* **61**, 3470–3478.
- Bermingham, A. & Collins, P. L. (1999).** The M2-2 protein of human respiratory syncytial virus is a regulatory factor involved in the balance between RNA replication and transcription. *Proc Natl Acad Sci USA* **96**, 11259–11264.
- Bhella, D., Ralph, A., Murphy, L. B. & Yeo, R. P. (2002).** Significant differences in nucleocapsid morphology within the *Paramyxoviridae*. *J Gen Virol* **83**, 1831–1839.
- Brown, G., Rixon, H. W. McL., Steel, J., McDonald, T. P., Pitt, A. R., Graham, S. & Sugrue, R. J. (2005).** Evidence for an association between heat shock protein 70 and the respiratory syncytial virus polymerase complex within lipid-raft membranes during virus infection. *Virology* **338**, 69–80.
- Browning, D. F. & Busby, S. J. W. (2004).** The regulation of bacterial transcription initiation. *Nat Rev Microbiol* **2**, 57–65.
- Bukreyev, A., Murphy, B. R. & Collins, P. L. (2000).** Respiratory syncytial virus can tolerate an intergenic sequence of at least 160 nucleotides with little effect on transcription or replication in vitro and in vivo. *J Virol* **74**, 11017–11026.
- Burke, E., Dupuy, L., Wall, C. & Barik, S. (1998).** Role of cellular actin in the gene expression and morphogenesis of human respiratory syncytial virus. *Virology* **252**, 137–148.
- Burke, E., Mahoney, N. M., Almo, S. C. & Barik, S. (2000).** Profilin is required for optimal actin-dependent transcription of respiratory syncytial virus genome RNA. *J Virol* **74**, 669–675.
- Calain, P. & Roux, L. (1993).** The rule of six, a basic feature for efficient replication of Sendai virus defective interfering RNA. *J Virol* **67**, 4822–4830.
- Cartee, T. L. & Wertz, G. W. (2001).** Respiratory syncytial virus M2-1 protein requires phosphorylation for efficient function and binds viral RNA during infection. *J Virol* **75**, 12188–12197.

Cartee, T. L., Megaw, A. G., Oomens, A. G. P. & Wertz, G. W. (2003). Identification of a single amino acid change in the human respiratory syncytial virus L protein that affects transcriptional termination. *J Virol* **77**, 7352–7360.

Castagné, N., Barbier, A., Bernard, J., Rezaei, H., Huet, J.-C., Henry, C., Da Costa, B. & Eléouët, J.-F. (2004). Biochemical characterization of the respiratory syncytial virus P–P and P–N complexes and localization of the P protein oligomerization domain. *J Gen Virol* **85**, 1643–1653.

Cheng, X., Park, H., Zhou, H. & Jin, H. (2005). Overexpression of the M2-2 protein of respiratory syncytial virus inhibits viral replication. *J Virol* **79**, 13943–13952.

Chuang, J. L. & Perrault, J. (1997). Initiation of vesicular stomatitis virus mutant polR1 transcription internally at the N gene in vitro. *J Virol* **71**, 1466–1475.

Collins, P. L. & Murphy, B. R. (2002). Respiratory syncytial virus: reverse genetics and vaccine strategies. *Virology* **296**, 204–211.

Collins, P. L. & Wertz, G. W. (1983). cDNA cloning and transcriptional mapping of nine polyadenylated RNAs encoded by the genome of human respiratory syncytial virus. *Proc Natl Acad Sci U S A* **80**, 3208–3212.

Collins, P. L. & Wertz, G. W. (1985). Nucleotide sequences of the 1B and 1C nonstructural protein mRNAs of human respiratory syncytial virus. *Virology* **143**, 442–451.

Collins, P. L., Dickens, L. E., Buckler-White, A., Olmsted, R. A., Spriggs, M. K., Camargo, E. & Coelingh, K. V. W. (1986). Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order. *Proc Natl Acad Sci U S A* **83**, 4594–4598.

Collins, P. L., Olmsted, R. A., Spriggs, M. K., Johnson, P. R. & Buckler-White, A. J. (1987). Gene overlap and site-specific attenuation of transcription of the viral polymerase L gene of human respiratory syncytial virus. *Proc Natl Acad Sci U S A* **84**, 5134–5138.

Collins, P. L., Mink, M. A. & Stec, D. S. (1991). Rescue of synthetic analogs of respiratory syncytial virus genomic RNA and effect of truncations and mutations on the expression of a foreign reporter gene. *Proc Natl Acad Sci U S A* **88**, 9663–9667.

Collins, P. L., Hill, M. G., Cristina, J. & Grosfeld, H. (1996). Transcription elongation factor of respiratory syncytial virus, a nonsegmented negative-strand RNA virus. *Proc Natl Acad Sci U S A* **93**, 81–85.

Collins, P. L., Whitehead, S. S., Bukreyev, A., Fearn, R., Teng, M. N., Juhasz, K., Chanock, R. M. & Murphy, B. R. (1999). Rational design of live-attenuated recombinant vaccine virus for human respiratory syncytial virus by reverse genetics. *Adv Virus Res* **54**, 423–451.

Collins, P. L., Chanock, R. M. & Murphy, B. R. (2001). Respiratory syncytial virus. In *Fields Virology*, 4th edn, pp. 1443–1485. Edited by D. M. Knipe & P. M. Howley. Philadelphia, PA: Lippincott Williams & Wilkins.

Cowton, V. M. & Fearn, R. (2005). Evidence that the respiratory syncytial virus polymerase is recruited to nucleotides 1 to 11 at the 3' end of the nucleocapsid and can scan to access internal signals. *J Virol* **79**, 11311–11322.

Cuesta, I., Geng, X., Asenjo, A. & Villanueva, N. (2000). Structural phosphoprotein M2-1 of the human respiratory syncytial virus is an RNA binding protein. *J Virol* **74**, 9858–9867.

Dickens, L. E., Collins, P. L. & Wertz, G. W. (1984). Transcriptional mapping of human respiratory syncytial virus. *J Virol* **52**, 364–369.

Dupuy, L. C., Dobson, S., Bitko, V. & Barik, S. (1999). Casein kinase 2-mediated phosphorylation of respiratory syncytial virus phosphoprotein P is essential for the transcription elongation activity of the viral polymerase; phosphorylation by casein kinase 1 occurs mainly at Ser²¹⁵ and is without effect. *J Virol* **73**, 8384–8392.

Dvir, A. (2002). Promoter escape by RNA polymerase II. *Biochim Biophys Acta* **1577**, 208–223.

Fearns, R. & Collins, P. L. (1999a). Model for polymerase access to the overlapped L gene of respiratory syncytial virus. *J Virol* **73**, 388–397.

Fearns, R. & Collins, P. L. (1999b). Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription. *J Virol* **73**, 5852–5864.

Fearns, R., Peeples, M. E. & Collins, P. L. (1997). Increased expression of the N protein of respiratory syncytial virus stimulates minigenome replication but does not alter the balance between the synthesis of mRNA and antigenome. *Virology* **236**, 188–201.

Fearns, R., Collins, P. L. & Peeples, M. E. (2000). Functional analysis of the genomic and antigenomic promoters of human respiratory syncytial virus. *J Virol* **74**, 6006–6014.

Fearns, R., Peeples, M. E. & Collins, P. L. (2002). Mapping the transcription and replication promoters of respiratory syncytial virus. *J Virol* **76**, 1663–1672.

Ferron, F., Longhi, S., Henrissat, B. & Canard, B. (2002). Viral RNA-polymerases – a predicted 2'-O-ribose methyltransferase domain shared by all *Mononegavirales*. *Trends Biochem Sci* **27**, 222–224.

Finke, S. & Conzelmann, K.-K. (1997). Ambisense gene expression from recombinant rabies virus: random packaging of positive- and negative-strand ribonucleoprotein complexes into rabies virions. *J Virol* **71**, 7281–7288.

García, J., García-Barreno, B., Vivo, A. & Melero, J. A. (1993). Cytoplasmic inclusions of respiratory syncytial virus-infected cells: formation of inclusion bodies in transfected cells that coexpress the nucleoprotein, the phosphoprotein, and the 22K protein. *Virology* **195**, 243–247.

García-Barreno, B., Jorcano, J. L., Aukenbauer, T., Lopez-Galindez, C. & Melero, J. A. (1988). Participation of cytoskeletal intermediate filaments in the infectious cycle of human respiratory syncytial virus (RSV). *Virus Res* **9**, 307–321.

García-Barreno, B., Delgado, T. & Melero, J. A. (1996). Identification of protein regions involved in the interaction of human respiratory syncytial virus phosphoprotein and nucleoprotein: significance for nucleocapsid assembly and formation of cytoplasmic inclusions. *J Virol* **70**, 801–808.

García-Barreno, B., Steel, J., Payá, M., Martínez-Sobrido, L., Delgado, T., Yeo, R. P. & Melero, J. A. (2005). Epitope mapping of human respiratory syncytial virus 22K transcription antitermination factor: role of N-terminal sequences in protein folding. *J Gen Virol* **86**, 1103–1107.

Grzelishvili, V. Z., Smallwood, S., Tower, D., Hall, R. L., Hunt, D. M. & Moyer, S. A. (2005). A single amino acid change in the L-polymerase protein of vesicular stomatitis virus completely abolishes viral mRNA cap methylation. *J Virol* **79**, 7327–7337.

Greive, S. J. & von Hippel, P. H. (2005). Thinking quantitatively about transcriptional regulation. *Nat Rev Mol Cell Biol* **6**, 221–232.

Grosfeld, H., Hill, M. G. & Collins, P. L. (1995). RNA replication by respiratory syncytial virus (RSV) is directed by the N, P, and L proteins; transcription also occurs under these conditions but requires RSV superinfection for efficient synthesis of full-length mRNA. *J Virol* **69**, 5677–5686.

Gubbay, O., Curran, J. & Kolakofsky, D. (2001). Sendai virus genome synthesis and assembly are coupled: a possible mechanism to promote viral RNA polymerase processivity. *J Gen Virol* **82**, 2895–2903.

Hardy, R. W. & Wertz, G. W. (1998). The product of the respiratory syncytial virus M2 gene ORF1 enhances readthrough of intergenic junctions during viral transcription. *J Virol* **72**, 520–526.

Hardy, R. W. & Wertz, G. W. (2000). The Cys₃-His₁ motif of the respiratory syncytial virus M2-1 protein is essential for protein function. *J Virol* **74**, 5880–5885.

Hardy, R. W., Harmon, S. B. & Wertz, G. W. (1999). Diverse gene junctions of respiratory syncytial virus modulate the efficiency of transcription termination and respond differently to M2-mediated antitermination. *J Virol* **73**, 170–176.

Harmon, S. B. & Wertz, G. W. (2002). Transcriptional termination modulated by nucleotides outside the characterized gene end sequence of respiratory syncytial virus. *Virology* **300**, 304–315.

Harmon, S. B., Megaw, A. G. & Wertz, G. W. (2001). RNA sequences involved in transcriptional termination of respiratory syncytial virus. *J Virol* **75**, 36–44.

Hinzman, E. E., Barr, J. N. & Wertz, G. W. (2002). Identification of an upstream sequence element required for vesicular stomatitis virus mRNA transcription. *J Virol* **76**, 7632–7641.

Horikami, S. M., Curran, J., Kolakofsky, D. & Moyer, S. A. (1992). Complexes of Sendai virus NP-P and P-L proteins are required for defective interfering particle genome replication in vitro. *J Virol* **66**, 4901–4908.

Hsu, L. M. (2002). Promoter clearance and escape in prokaryotes. *Biochim Biophys Acta* **1577**, 191–207.

Isemi, F., Baudin, F., Blondel, D. & Ruigrok, R. W. (2000). Structure of the RNA inside the vesicular stomatitis virus nucleocapsid. *RNA* **6**, 270–281.

Isemi, F., Baudin, F., Garcin, D., Marq, J. B., Ruigrok, R. W. & Kolakofsky, D. (2002). Chemical modification of nucleotide bases and mRNA editing depend on hexamer or nucleoprotein phase in Sendai virus nucleocapsids. *RNA* **8**, 1056–1067.

Iverson, L. E. & Rose, J. K. (1981). Localized attenuation and discontinuous synthesis during vesicular stomatitis virus transcription. *Cell* **23**, 477–484.

Jin, H., Cheng, X., Zhou, H. Z. Y., Li, S. & Seddiqui, A. (2000). Respiratory syncytial virus that lacks open reading frame 2 of the M2 gene (M2-2) has altered growth characteristics and is attenuated in rodents. *J Virol* **74**, 74–82.

Johnson, P. R. & Collins, P. L. (1988). The A and B subgroups of human respiratory syncytial virus: comparison of intergenic and gene-overlap sequences. *J Gen Virol* **69**, 2901–2906.

Juhász, K., Murphy, B. R. & Collins, P. L. (1999). The major attenuating mutations of the respiratory syncytial virus vaccine candidate *cpts530/1009* specify temperature-sensitive defects in transcription and replication and a non-temperature-sensitive alteration in mRNA termination. *J Virol* **73**, 5176–5180.

Kallewaard, N. L., Bowen, A. L. & Crowe, J. E., Jr (2005). Cooperativity of actin and microtubule elements during replication of respiratory syncytial virus. *Virology* **331**, 73–81.

Karlin, D., Ferron, F., Canard, B. & Longhi, S. (2003). Structural disorder and modular organization in *Paramyxovirinae* N and P. *J Gen Virol* **84**, 3239–3252.

Khattar, S. K., Yunus, A. S., Collins, P. L. & Samal, S. K. (2000). Mutational analysis of the bovine respiratory syncytial virus nucleocapsid protein using a minigenome system: mutations that affect encapsidation, RNA synthesis, and interaction with the phosphoprotein. *Virology* **270**, 215–228.

Kingston, R. L., Hamel, D. J., Gay, L. S., Dahlquist, F. W. & Matthews, B. W. (2004). Structural basis for the attachment of a paramyxoviral polymerase to its template. *Proc Natl Acad Sci U S A* **101**, 8301–8306.

Kolakofsky, D., Pelet, T., Garcin, D., Hausmann, S., Curran, J. & Roux, L. (1998). Paramyxovirus RNA synthesis and the requirement for hexamer genome length: the rule of six revisited. *J Virol* **72**, 891–899.

Kolakofsky, D., Le Mercier, P., Iseni, F. & Garcin, D. (2004). Viral RNA polymerase scanning and the gymnastics of Sendai virus RNA synthesis. *Virology* **318**, 463–473.

Krempl, C., Murphy, B. R. & Collins, P. L. (2002). Recombinant respiratory syncytial virus with the G and F genes shifted to the promoter-proximal positions. *J Virol* **76**, 11931–11942.

Kuo, L., Grosfeld, H., Cristina, J., Hill, M. G. & Collins, P. L. (1996a). Effects of mutations in the gene-start and gene-end sequence motifs on transcription of monocistronic and dicistronic minigenomes of respiratory syncytial virus. *J Virol* **70**, 6892–6901.

Kuo, L., Fearn, R. & Collins, P. L. (1996b). The structurally diverse intergenic regions of respiratory syncytial virus do not modulate sequential transcription by a dicistronic minigenome. *J Virol* **70**, 6143–6150.

Kuo, L., Fearn, R. & Collins, P. L. (1997). Analysis of the gene start and gene end signals of human respiratory syncytial virus: quasi-templated initiation at position 1 of the encoded mRNA. *J Virol* **71**, 4944–4953.

Le Mercier, P., Garcin, D., Hausmann, S. & Kolakofsky, D. (2002). Ambisense Sendai viruses are inherently unstable but are useful to study viral RNA synthesis. *J Virol* **76**, 5492–5502.

Le Mercier, P., Garcin, D., Garcia, E. & Kolakofsky, D. (2003). Competition between the Sendai virus N mRNA start site and the genome 3'-end promoter for viral RNA polymerase. *J Virol* **77**, 9147–9155.

Li, T. & Pattnaik, A. K. (1999). Overlapping signals for transcription and replication at the 3' terminus of the vesicular stomatitis virus genome. *J Virol* **73**, 444–452.

Li, J., Fontaine-Rodriguez, E. C. & Whelan, S. P. J. (2005). Amino acid residues within conserved domain VI of the vesicular stomatitis virus large polymerase protein essential for RNA cap methyltransferase activity. *J Virol* **79**, 13373–13384.

Liuzzi, M., Mason, S. W., Cartier, M. & 17 other authors (2005). Inhibitors of respiratory syncytial virus replication target cotranscriptional mRNA guanylation by viral RNA-dependent RNA polymerase. *J Virol* **79**, 13105–13115.

Lu, B., Ma, C.-H., Brazas, R. & Jin, H. (2002). The major phosphorylation sites of the respiratory syncytial virus phosphoprotein are dispensable for virus replication in vitro. *J Virol* **76**, 10776–10784.

Mason, S. W., Aberg, E., Lawetz, C., DeLong, R., Whitehead, P. & Liuzzi, M. (2003).

Interaction between human respiratory syncytial virus (RSV) M2-1 and P proteins is required for reconstitution of M2-1-dependent RSV minigenome activity. *J Virol* **77**, 10670–10676.

Mason, S. W., Lawetz, C., Gaudette, Y., Dô, F., Scouten, E., Lagacé, L., Simoneau, B. & Liuzzi, M. (2004). Polyadenylation-dependent screening assay for respiratory syncytial virus RNA transcriptase activity and identification of an inhibitor. *Nucleic Acids Res* **32**, 4758–4767.

Mazumder, B. & Barik, S. (1994). Requirement of casein kinase II-mediated phosphorylation for the transcriptional activity of human respiratory syncytial viral phosphoprotein P: transdominant negative phenotype of phosphorylation-defective P mutants. *Virology* **205**, 104–111.

McDonald, T. P., Pitt, A. R., Brown, G., Rixon, H. W. McL. & Sugrue, R. J. (2004). Evidence that the respiratory syncytial virus polymerase complex associates with lipid rafts in virus-infected cells: a proteomic analysis. *Virology* **330**, 147–157.

McGivern, D. R., Collins, P. L. & Fearn, R. (2005). Identification of internal sequences in the 3' leader region of human respiratory syncytial virus that enhance transcription and confer replication processivity. *J Virol* **79**, 2449–2460.

Mellon, M. G. & Emerson, S. U. (1978). Rebinding of transcriptase components (L and NS proteins) to the nucleocapsid template of vesicular stomatitis virus. *J Virol* **27**, 560–567.

Mink, M. A., Stec, D. S. & Collins, P. L. (1991). Nucleotide sequences of the 3' leader and 5' trailer regions of human respiratory syncytial virus genomic RNA. *Virology* **185**, 615–624.

Moudy, R. M., Harmon, S. B., Sullender, W. M. & Wertz, G. W. (2003). Variations in transcription termination signals of human respiratory syncytial virus clinical isolates affect gene expression. *Virology* **313**, 250–260.

Moudy, R. M., Sullender, W. M. & Wertz, G. W. (2004). Variations in intergenic region sequences of human respiratory syncytial virus clinical isolates: analysis of effects on transcriptional regulation. *Virology* **327**, 121–133.

Müller, R., Poch, O., Delarue, M., Bishop, D. H. L. & Bouloy, M. (1994). Rift valley fever virus L segment: correction of the sequence and possible functional role of newly identified regions conserved in RNA-dependent polymerases. *J Gen Virol* **75**, 1345–1352.

Murphy, S. K. & Parks, G. D. (1999). RNA replication for the paramyxovirus simian virus 5 requires an internal repeated (CGNNNN) sequence motif. *J Virol* **73**, 805–809.

Murphy, S. K., Ito, Y. & Parks, G. D. (1998). A functional antigenomic promoter for the paramyxovirus simian virus 5 requires proper spacing between an essential internal segment and the 3' terminus. *J Virol* **72**, 10–19.

Murphy, L. B., Loney, C., Murray, J., Bhella, D., Ashton, P. & Yeo, R. P. (2003). Investigations into the amino-terminal domain of the respiratory syncytial virus nucleocapsid protein reveal elements important for nucleocapsid formation and interaction with the phosphoprotein. *Virology* **307**, 143–153.

Ogino, T., Kobayashi, M., Iwama, M. & Mizumoto, K. (2005). Sendai virus RNA-dependent RNA polymerase L protein catalyzes cap methylation of virus-specific mRNA. *J Biol Chem* **280**, 4429–4435.

O'Reilly, E. K. & Kao, C. C. (1998). Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure. *Virology* **252**, 287–303.

Orphanides, G. & Reinberg, D. (2002). A unified theory of gene expression. *Cell* **108**, 439–451.

Patton, J. T., Davis, N. L. & Wertz, G. W. (1984). N protein alone satisfies the requirement for protein synthesis during RNA replication of vesicular stomatitis virus. *J Virol* **49**, 303–309.

Peeples, M. E. & Collins, P. L. (2000). Mutations in the 5' trailer region of a respiratory syncytial virus minigenome which limit RNA replication to one step. *J Virol* **74**, 146–155.

Perrault, J. & McLearn, P. W. (1984). ATP dependence of vesicular stomatitis virus transcription initiation and modulation by mutation in the nucleocapsid protein. *J Virol* **51**, 635–642.

Poch, O., Sauvaget, I., Delarue, M. & Tordo, N. (1989). Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J* **8**, 3867–3874.

Poch, O., Blumberg, B. M., Bougueleret, L. & Tordo, N. (1990). Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. *J Gen Virol* **71**, 1153–1162.

Qanungo, K. R., Shaji, D., Mathur, M. & Banerjee, A. K. (2004). Two RNA polymerase complexes from vesicular stomatitis virus-infected cells that carry out transcription and replication of genome RNA. *Proc Natl Acad Sci U S A* **101**, 5952–5957.

Rose, J. K. (1975). Heterogeneous 5'-terminal structures occur on vesicular stomatitis virus mRNAs. *J Biol Chem* **250**, 8098–8104.

Rose, J. K., Lodish, H. F. & Brock, M. L. (1977). Giant heterogeneous polyadenylic acid on vesicular stomatitis virus mRNA synthesized in vitro in the presence of S-adenosylhomocysteine. *J Virol* **21**, 683–693.

Samal, S. K. & Collins, P. L. (1996). RNA replication by a respiratory syncytial virus RNA analog does not obey the rule of six and retains a nonviral trinucleotide extension at the leader end. *J Virol* **70**, 5075–5082.

Sánchez-Seco, M. P., Navarro, J., Martínez, R. & Villanueva, N. (1995). C-Terminal phosphorylation of human respiratory syncytial virus P protein occurs mainly at serine residue 232. *J Gen Virol* **76**, 425–430.

Shilatifard, A. (2004). Transcriptional elongation control by RNA polymerase II: a new frontier. *Biochim Biophys Acta* **1677**, 79–86.

Stec, D. S., Hill, M. G., III & Collins, P. L. (1991). Sequence analysis of the polymerase L gene of human respiratory syncytial virus and predicted phylogeny of nonsegmented negative-strand viruses. *Virology* **183**, 273–287.

Stillman, E. A. & Whitt, M. A. (1999). Transcript initiation and 5'-end modification are separable events during vesicular stomatitis virus transcription. *J Virol* **73**, 7199–7209.

Sun, J.-H. & Kao, C. C. (1997a). RNA synthesis by the brome mosaic virus RNA-dependent RNA polymerase: transition from initiation to elongation. *Virology* **233**, 63–73.

Sun, J.-H. & Kao, C. C. (1997b). Characterization of RNA products associated with or aborted by a viral RNA-dependent RNA polymerase. *Virology* **236**, 348–353.

Sun, J.-H., Adkins, S., Faurote, G. & Kao, C. C. (1996). Initiation of (–)-strand RNA synthesis catalyzed by the BMV RNA-dependent RNA polymerase: synthesis of oligonucleotides. *Virology* **226**, 1–12.

Sutherland, K. A., Collins, P. L. & Peeples, M. E. (2001). Synergistic effects of gene-end signal mutations and the M2-1 protein on transcription termination by respiratory syncytial virus. *Virology* **288**, 295–307.

Tang, R. S., Nguyen, N., Cheng, X. & Jin, H. (2001). Requirement of cysteines and length of the human respiratory syncytial virus M2-1 protein for protein function and virus viability. *J Virol* **75**, 11328–11335.

Tapparel, C., Maurice, D. & Roux, L. (1998). The activity of Sendai virus genomic and antigenomic promoters requires a second element past the leader template regions: a motif (GNNNNN)₃ is essential for replication. *J Virol* **72**, 3117–3128.

Teng, M. N. & Collins, P. L. (1998). Identification of the respiratory syncytial virus proteins required for formation and passage of helper-dependent infectious particles. *J Virol* **72**, 5707–5716.

Testa, D. & Banerjee, A. K. (1979). Initiation of RNA synthesis *in vitro* by vesicular stomatitis virus: role of ATP. *J Biol Chem* **254**, 2053–2058.

Tolley, K. P., Marriott, A. C., Simpson, A. & 9 other authors (1996). Identification of mutations contributing to the reduced virulence of a modified strain of respiratory syncytial virus. *Vaccine* **14**, 1637–1646.

Tran, K. C., Collins, P. L. & Teng, M. N. (2004). Effects of altering the transcription termination signals of respiratory syncytial virus on viral gene expression and growth *in vitro* and *in vivo*. *J Virol* **78**, 692–699.

Ulloa, L., Serra, R., Asenjo, A. & Villanueva, N. (1998). Interactions between cellular actin and human respiratory syncytial virus (HRSV). *Virus Res* **53**, 13–25.

Vidal, S. & Kolakofsky, D. (1989). Modified model for the switch from Sendai virus transcription to replication. *J Virol* **63**, 1951–1958.

Villanueva, N., Hardy, R., Asenjo, A., Yu, Q. & Wertz, G. (2000). The bulk of the phosphorylation of human respiratory syncytial virus phosphoprotein is not essential but modulates viral RNA transcription and replication. *J Gen Virol* **81**, 129–133.

von Hippel, P. H. (2004). Completing the view of transcriptional regulation. *Science* **305**, 350–352.

Vulliémoz, D. & Roux, L. (2001). “Rule of six”: how does the Sendai virus RNA polymerase keep count? *J Virol* **75**, 4506–4518.

Vulliémoz, D. & Roux, L. (2002). Given the opportunity, the Sendai virus RNA-dependent RNA polymerase could as well enter its template internally. *J Virol* **76**, 7987–7995.

Vulliémoz, D., Cordey, S., Mottet-Osman, G. & Roux, L. (2005). Nature of a paramyxovirus replication promoter influences a nearby transcription signal. *J Gen Virol* **86**, 171–180.

Whelan, S. P. J. & Wertz, G. W. (1999). Regulation of RNA synthesis by the genomic termini of vesicular stomatitis virus: identification of distinct sequences essential for transcription but not replication. *J Virol* **73**, 297–306.

Whelan, S. P. J. & Wertz, G. W. (2002). Transcription and replication initiate at separate sites on the vesicular stomatitis virus genome. *Proc Natl Acad Sci U S A* **99**, 9178–9183.

Whelan, S. P. J., Barr, J. N. & Wertz, G. W. (2000). Identification of a minimal size requirement for termination of vesicular stomatitis virus mRNA: implications for the mechanism of transcription. *J Virol* **74**, 8268–8276.

- Whelan, S. P. J., Barr, J. N. & Wertz, G. W. (2004).** Transcription and replication of nonsegmented negative-strand RNA viruses. *Curr Top Microbiol Immunol* **283**, 61–119.
- Young, B. A., Gruber, T. M. & Gross, C. A. (2002).** Views of transcription initiation. *Cell* **109**, 417–420.
- Yu, Q., Hardy, R. W. & Wertz, G. W. (1995).** Functional cDNA clones of the human respiratory syncytial (RS) virus N, P, and L proteins support replication of RS virus genomic RNA analogs and define minimal *trans*-acting requirements for RNA replication. *J Virol* **69**, 2412–2419.
- Zorio, D. A. R. & Bentley, D. L. (2004).** The link between mRNA processing and transcription: communication works both ways. *Exp Cell Res* **296**, 91–97.

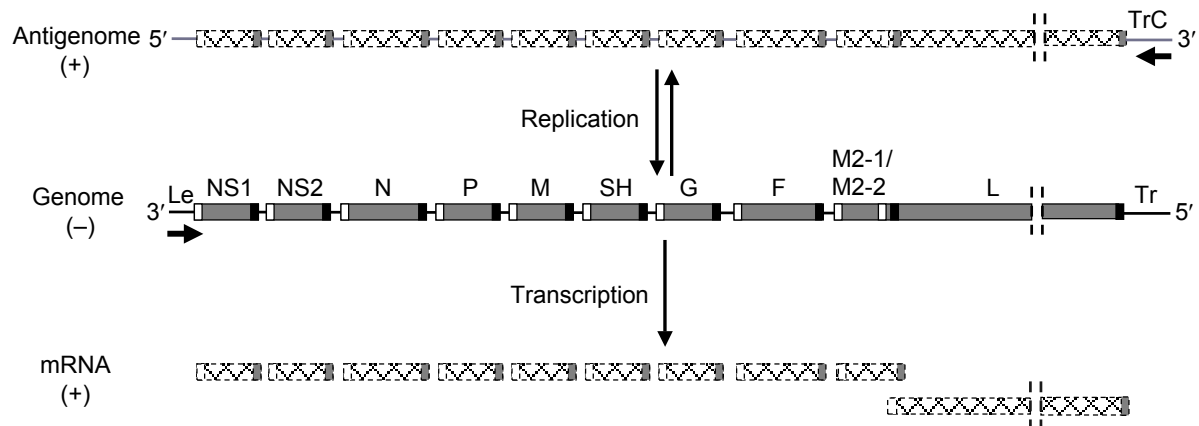


Fig. 1. Schematic diagram (not to scale) depicting the RSV genome and its transcription and replication products. The virus genes are depicted as grey rectangles; the L gene, which comprises almost half of the genome, has been truncated. The GS and GE signals are shown as white and black boxes, respectively. The encoded antigenome and mRNAs are indicated by hatched rectangles. Arrows indicate the location of the promoters.

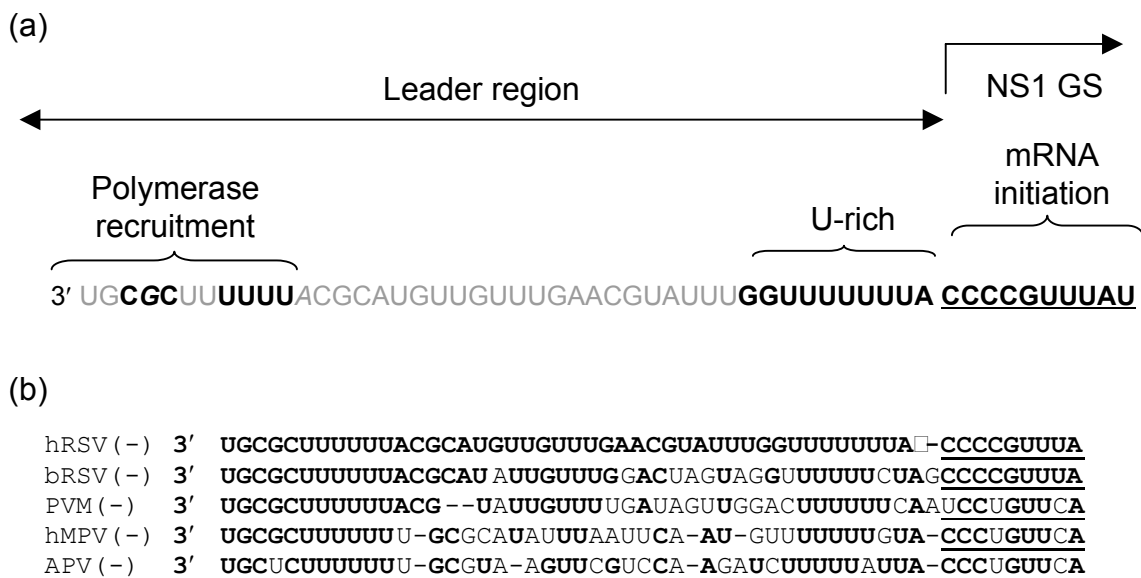


Fig. 2. *cis*-acting sequences involved in transcription initiation. (a) The RSV Le and GS sequence, with the regions required for transcription shown in bold type and the nucleotides not required shown in normal type. The GS signal for the first gene (NS1) is underlined. Substitutions at positions 4 and 12 (italicized) can inhibit transcription and augment replication, indicating that these are optimized for transcription. (b) Alignment of the 3' termini of the genomes of members of the *Pneumovirinae*: RSV A2 (hRSV), bovine RSV (bRSV), pneumonia virus of mice (PVM), human metapneumovirus (hMPV) and avian metapneumovirus (APV). GenBank accession numbers are AF035006, NC_001989, AY729016, NC_004148 and AY590688, respectively. The sequence is shown as negative-sense RNA and nucleotides identical to the RSV genome sequence are highlighted in bold type. Gaps, indicated by dashes, were introduced to maximize alignment. The GS sequences are underlined.

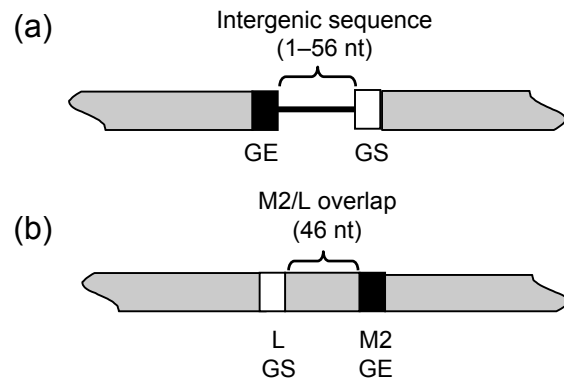


Fig. 3. Diagram showing the arrangement of signals at the typical and M2/L gene junctions. The virus genes are depicted as grey rectangles and the GS and GE sequences are shown as white and black boxes, respectively. (a) Arrangement of a typical gene junction, in which the intergenic region can vary from 1 to 56 nt. (b) M2/L gene junction, in which there are 46 nt between the end of the L GS signal and the start of the M2 GE signal.

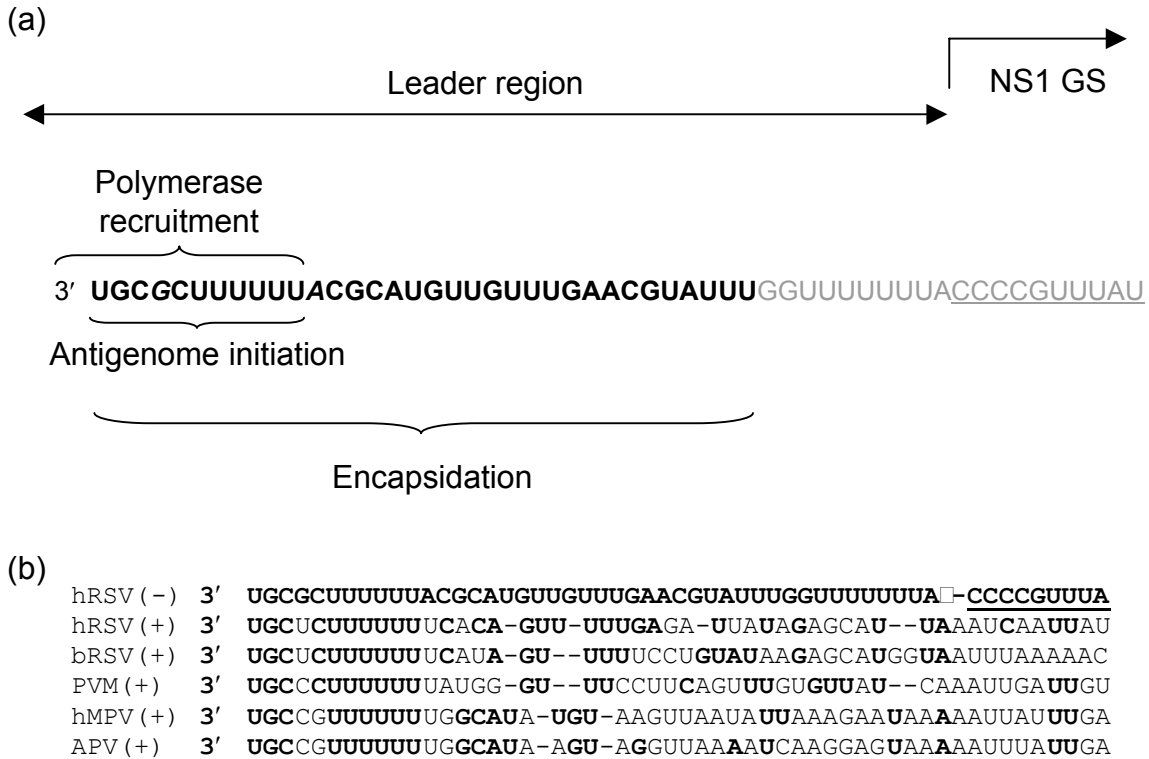


Fig. 4. *cis*-acting sequences involved in replication initiation. (a) The RSV Le and GS sequence, with the regions required for replication shown in bold type and the nucleotides that are not required shown in normal type. Note that the nucleotides required for encapsidation have not been mapped precisely. The GS signal is underlined. (b) Alignment of the 3' termini of the antigenome RNAs of members of the *Pneumovirinae*. The sequences are shown aligned to the 3' terminus of the RSV genome, with identical nucleotides in bold type. Spaces introduced to maximize alignment are indicated by dashes. The GS sequence in the RSV genome is underlined. Abbreviations and GenBank accession numbers are the same as for Fig. 2.

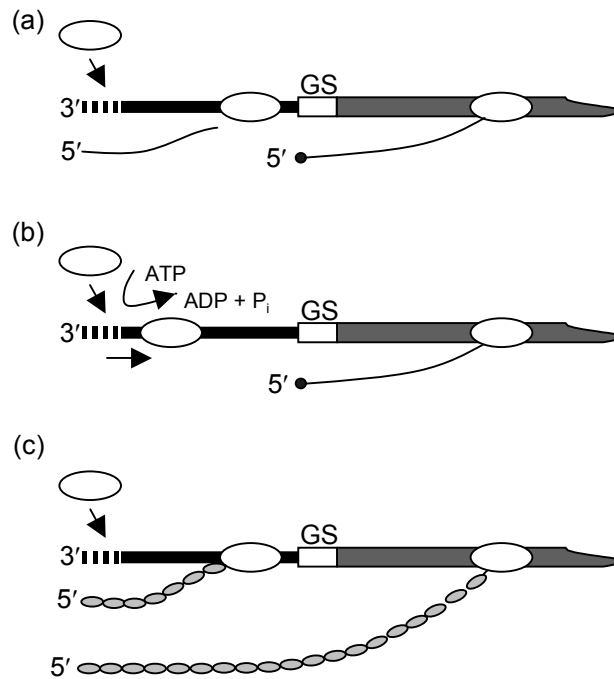


Fig. 5. Diagram depicting the models for transcription and replication initiation from the Le promoter region. The Le region is shown as a thick black line and the polymerase-binding sequence (nt 1–11) is hatched. The NS1 gene sequence is in dark grey and its GS signal is represented by a white box (□). The polymerase active site is represented by an oval and the 5' cap on the mRNA is shown as a filled circle (●). N protein coating the antigenome is shown in light grey. (a, b) Transcription models 1 and 2, respectively. (c) A model for replication.