Virology 273, 1–8 (2000)
doi:10.1006/viro.2000.0421, available online at http://www.idealibrary.com on IDEAL®

MINIREVIEW

Synthesis of Subgenomic RNAs by Positive-Strand RNA Viruses

W. Allen Miller[1] and Gennadiy Koev[2]

Plant Pathology Department, Iowa State University, Ames, Iowa 50011

Received March 1, 2000; returned to author for revision March 31, 2000; accepted May 11, 2000

Many RNA viruses encode more than one gene on a single genomic RNA. Yet only the first gene, or open reading frame (ORF), on a normal eukaryotic mRNA is translated. Thus, downstream genes on viral genomes are expressed either via novel translational events or, more commonly, by deployment of subgenomic mRNAs (sgRNAs). Subgenomic RNAs of positive-strand viruses have the same 3′ ends as genomic RNA, but have deletions at the 5′ ends to bring the 5′ end of the RNA in proximity with the start codon of downstream (on genomic RNA) ORFs. Because replication is required for sgRNA synthesis, the RNA-dependent RNA polymerase (RdRp) is always translated first, directly from genomic RNA of positive-strand RNA viruses. sgRNAs express products needed during intermediate and late stages of infection, such as structural or movement proteins.

Taxa of (+) sense viruses that produce sgRNAs include animal viruses in the order *Nidovirales* (*Coronaviridae* and *Arteriviridae* families) and the families *Togaviridae, Caliciviridae, Nodaviridae,* and *Astroviridae,* plant viruses of the *Luteoviridae, Bromoviridae, Tombusviridae,* and *Closteroviridae* families, and the *Tobravirus, Carlavirus, Tymovirus, Potexvirus, Hordeivirus, Tobamovirus, Sobemovirus,* and *Furovirus* genera. Subgenomic RNA synthesis has been studied more in plant viruses than in animal viruses. This is probably because a greater percentage of all plant viruses make sgRNAs and also because plant viruses, with smaller genomes and highly efficient replication, are often more amenable than animal viruses to studies of RNA replication mechanisms, especially in cell-free extracts. The similarities between *Bromoviridae* and alphaviruses (below) provide one example of how plant viruses can serve as useful models for understanding replication of the larger animal virus RNAs.

Understanding mechanisms of sgRNA synthesis is important because it will shed light on how these RNA viruses replicate. Knowing how the decision between sgRNA synthesis and genomic RNA synthesis is made may allow us to interfere with both replication and gene expression of noxious viruses. Subgenomic RNA promoters may be recombination hot spots (Miller *et al.,* 1997; van Marle *et al.,* 1999a). A better understanding of the mechanism of their synthesis may provide insight into the origin and diversification of viruses by recombination. sgRNA-producing viruses are becoming used increasingly for expression of vaccines and other pharmaceutically useful proteins in plants (McCormick *et al.,* 1999) and in insect (Hahn *et al.,* 1992) and animal (van Dinten *et al.,* 1997) cells. Presence of a host sequence in a viral subgenomic RNA can induce rapid silencing of the host gene expression in plants (Baulcombe, 1999) and insect cells (Johnson *et al.,* 1999). This is being exploited for high-throughput functional genomics (Baulcombe, 1999; Kumagai *et al.,* 1995). To optimize these expression or silencing vectors, the mechanism of sgRNA synthesis must be understood.

Several basic mechanisms, for generating less-than-full-length viral RNAs from genomic RNA template, have been proposed (Fig. 1). The first is internal initiation, in which the replicase initiates (+) strand sgRNA transcription internally on a (−) strand copy of genomic RNA (Miller *et al.,* 1985). The second mechanism is premature termination during (−) strand synthesis from the genomic RNA template, giving a subgenomic-length (−) strand (Sit *et al.,* 1998). This would then serve as template for end-to-end (+) strand synthesis. Variations on the internal initiation and premature termination mechanisms involving discontinuous transcription have been proposed for *Nidovirales.* The 5′ end of each nidoviral sgRNA contains the sequence derived from the genomic

[1] To whom correspondence should be addressed at Plant Pathology Department and Center for Plant Responses to Environmental Stresses, 351 Bessey Hall, Iowa State University, Ames, IA 50011. Fax: 515-294-9420. E-mail: wamiller@iastate.edu.

[2] Current address: Howard Hughes Medical Institute, Department of Molecular Microbiology and Immunology, University of Southern California School of Medicine, 2011 Zonal Avenue, HMR-401, Los Angeles, CA 90033-1054. Fax: 323-342-9555. E-mail: koev@usc.edu.

0042-6822/00 $35.00

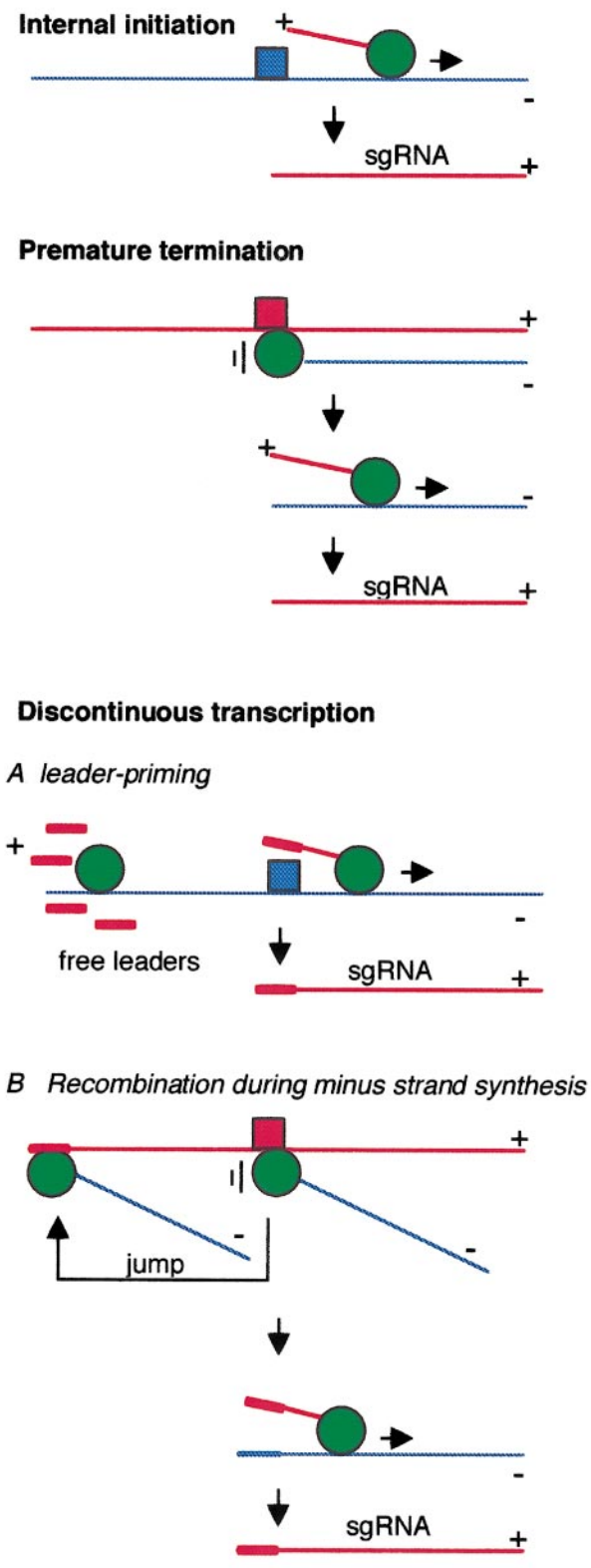

FIG. 1. Potential mechanisms of subgenomic RNA synthesis. Blue and red lines, (−) and (+) sense RNAs, respectively. Boxes, subgenomic RNA transcription initiation sites (blue) or termination sequences (red). Thick lines, 5′ genomic RNA leader (red) or its complement (blue). Green circles, viral replicases. Vertical bars, terminations of transcription.

RNA 5′ end. In this case, internal initiation may be primed by a short RNA derived from the 5′ end of the genome (Zhang and Lai, 1994). Alternatively, the sgRNAs may arise by replicase hopping from an internal termination site during (−) strand synthesis to the 5′ end of the genomic RNA template, without releasing nascent strand, and then run-off termination at the 5′ end of the genome (Sawicki and Sawicki, 1990). Other possibilities include cleavage and splicing of genomic RNA. These events are difficult to reconcile with maintenance of intact progeny viral genomes, and little evidence supports these possibilities.

For brevity, we refer to all sequences that give rise to sgRNAs as subgenomic RNA promoters. The initial mapping of subgenomic promoters of brome mosaic virus (BMV) (French and Ahlquist, 1988; Marsh *et al.,* 1988) and the alphavirus, Sindbis virus (Levis *et al.,* 1990), revealed a promoter of less than 100 bases, mostly upstream [in the (+) sense] of the transcription initiation site.

Other subgenomic promoters are more diverse and often more complex. They have been mapped in a variety of plant viruses (Balmori *et al.,* 1993; Boccard and Baulcombe, 1993; Haasnoot *et al.,* 2000; Johnston and Rochon, 1995; Koev *et al.,* 1999; van der Vossen *et al.,* 1995; Wang *et al.,* 1999; Wang and Simon, 1997; Zavriev *et al.,* 1996). They range from 24 nt (Wang and Simon, 1997) to over 100 nt in size (Balmori *et al.,* 1993; Koev and Miller, 2000; van der Vossen *et al.,* 1995). At least a portion of each promoter is located immediately upstream of the 5′ end of the resulting sgRNA (Fig. 2). However, essential components of some sgRNA promoters are located downstream of the 5′ end of the sgRNA, very distantly upstream, or even on a separate RNA molecule. The sgRNA promoter of beet necrotic yellow vein virus (BNYVV) (Balmori *et al.,* 1993) and two of the three sgRNA promoters of barley yellow dwarf virus (BYDV, Fig. 2) are located primarily downstream (Koev and Miller, 2000). Also, unlike the alpha-like virus group, sgRNA promoters from related viruses, or even within a virus, often have few or no features in common [e.g., tobacco mosaic virus (TMV) (Lehto *et al.,* 1990) and cucumber necrosis virus (CNV) (Johnston and Rochon, 1995)]. Other viruses contain short homologous oligonucleotides at the 5′ ends of their sgRNAs, with differing adjacent sequences, as reported in BNYVV (Balmori *et al.,* 1993) and tobacco necrosis virus (TNV) (Meulewaeter *et al.,* 1992). These differences between promoters may allow regulation of gene expression by differential transcription of sgRNAs. It also suggests that the protein components involved in regulating each promoter may differ (see below). The most extreme differences may be due to entirely different mechanisms for sgRNA synthesis.

The genomes of *Nidovirales* have particularly complex subgenomic RNA promoters. These viruses have very large RNA genomes (15 to over 31 kb) and produce a

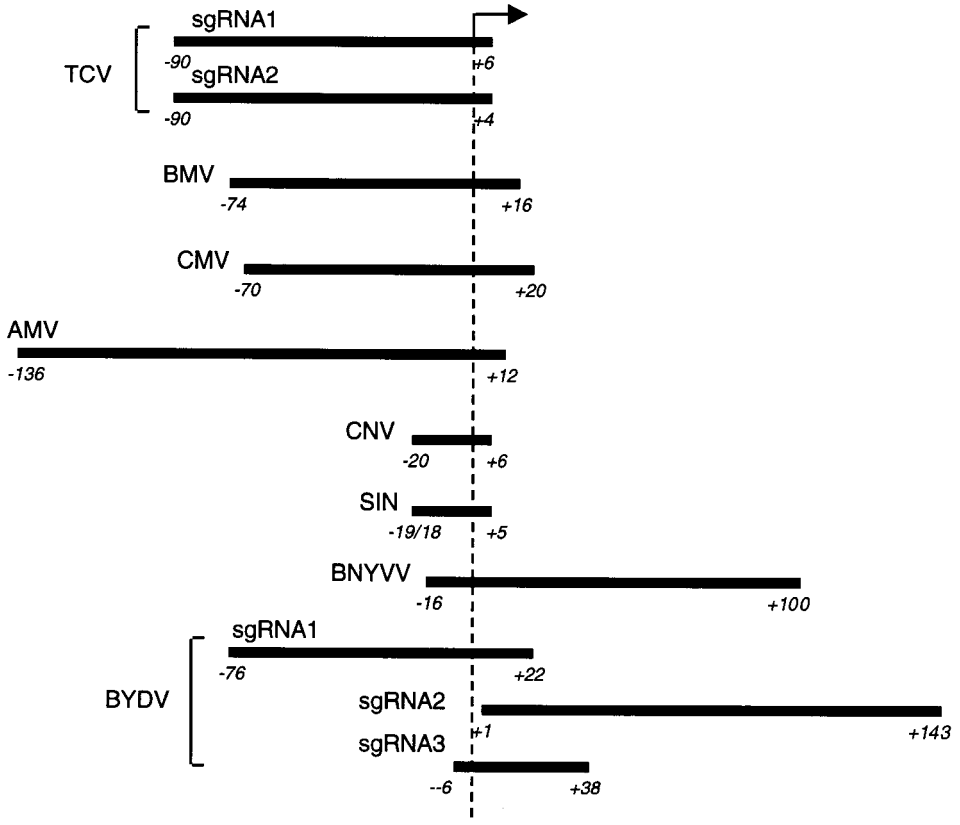

FIG. 2. Positions of subgenomic RNA promoters (continuous transcription) for which 5′ and 3′ extremities are known. Numbers indicate nucleotide positions relative to the sgRNA 5′ end (+1, indicated by dashed line). Abbreviations and references are in text.

nested set of seven 3′ coterminal sgRNAs (Lai and Cavanagh, 1997). All sgRNAs contain a 90-nt leader sequence derived from the 5′ end of the genomic RNA. This indicates that during transcription sgRNAs must somehow acquire the leader sequence from the 5′ end of gRNA [reviewed by Lai and Cavanagh (1997)]. A conserved motif found in the intergenic (IG) regions and in the 5′ leader, an adjacent sequence in the 5′ leader (Chang *et al.,* 1996), and distant sequences in the complement of the 3′ UTR (Lin *et al.,* 1996) all are necessary for mouse hepatitis coronavirus (MHV) sgRNA synthesis.

Other examples of distal *cis*-elements have been shown to regulate sgRNA synthesis, but in the absence of discontinuous transcription. Some of these elements exert their effect by long-distance base pairing to the regions near the start site. Mutations that disrupted potential base pairing between a region at the 5′ end of the potato virus X (PVX) genomic RNA (gRNA) and regions upstream of the transcription start sites of the sgRNAs abolished transcription. Compensatory mutations that restored base pairing also restored sgRNA transcription, although not to the wild-type level (Kim and Hemenway, 1996, 1999). This implies a requirement for both the base pairing and the primary RNA sequence of these elements. Also, a stable stem–loop structure at the 5′ end of PVX gRNA is important for both gRNA and sgRNA accumulation. Similar covariation mutagenesis of tomato bushy stunt tombusvirus (TBSV) RNA demonstrated a requirement for complementarity between a region immediately upstream of the sgRNA2 start site and a region located approximately 1000 nt upstream (Zhang *et al.,* 1999). In both PVX and TBSV, phylogenetic analysis showed conservation of such long-distance interactions in related viruses, reiterating their biological relevance.

Another interesting example of potential long-distance interactions in modulation of sgRNA transcription was observed in a TMV-based vector. Placement of three pseudoknots derived from the 3′ UTR of TMV behind a GFP ORF, in place of the viral coat protein, resulted in the increased transcription of the GFP sgRNA from the coat protein promoter (Shivprasad *et al.,* 1999). This also resulted in the decrease of transcription of the 3′ proximal sgRNA, suggesting that the pseudoknots redistributed viral polymerase activity localizing most of it to the nearest upstream subgenomic promoter.

The ultimate in long-distance interactions is in red clover necrotic mosaic virus (RCNMV), which has two genomic RNAs. Here, base pairing must occur between two different RNA molecules to form the sgRNA promoter. Sequence in the loop of a stem–loop in RCNMV genomic RNA2 must base pair to a sequence within RCNMV genomic RNA1 located just upstream of the

sgRNA 5′ end (Sit *et al.,* 1998). This is the only known example of a requirement for *trans*-activation of sgRNA transcription, i.e., a sgRNA promoter that is naturally split between two molecules. However, the leader-dependent mechanism of coronaviruses may occur by intermolecular recombination under some conditions (below).

Several mechanisms of sgRNA synthesis have been proposed (Fig. 1). The most widely recognized model is internal initiation of the replicase on longer-than-subgenomic-length (−) strand template. This mechanism applies to *Bromoviridae* (Miller *et al.,* 1985; van der Kuyl *et al.,* 1990) and other plant viruses (Wang and Simon, 1997) and the alphaviruses (Levis *et al.,* 1990). The subgenomic promoters described earlier and the RNA-dependent RNA polymerases (RdRp) of the *Bromoviridae* and alphaviruses are related as members of the alphavirus-like supergroup (Koonin and Dolja, 1993). Thus, RdRp itself may provide the template specificity, recognizing the conserved promoter elements (flanked by enhancer domains) in the (−) strand immediately upstream [in the (+) sense] of the base at which (+) strand synthesis initiates.

The above simple model clearly does not apply to all other RNA viruses, for example the *trans*-activation of transcription of RCNMV sgRNA (Sit *et al.,* 1998). The simplest model to explain this is that sgRNA is generated by premature termination of minus strand synthesis (Sit *et al.,* 1998). It was proposed that the base pairing between the two RNA sequences creates a termination structure, which alone, or with the help of protein factors, makes the viral RdRp pause and dissociate, releasing newly synthesized minus strand. This minus strand sgRNA would then serve as a template for plus strand sgRNA synthesis (Fig. 1). This was also proposed for TBSV sgRNA2, which requires cis base pairing between sequence elements more than a kilobase apart (Zhang *et al.,* 1999).

Requirement for secondary structure has been suggested to support a termination mechanism, because secondary structures facilitate termination in RNA recombination (Nagy *et al.,* 1998). However, probably all sgRNA promoters require secondary structure, including those known to be synthesized by internal initiation of transcription on the (−) strand (Wang and Simon, 1997; Haasnoot *et al.,* 2000). Conversely, rhabdovirus transcription termination depends only on primary sequence (Barr *et al.,* 1997), so termination can occur in the absence of secondary structure. Detection of the minus strand sgRNAs has been argued to support the premature termination model. However, minus strand RNA4 has been extracted from cells infected with BMV (Ishikawa *et al.,* 1997; Kroner *et al.,* 1990), even though BMV clearly employs internal initiation of sgRNA synthesis. The amount of double-stranded RNA4 is far less than the other three BMV RNAs, even though (+) strand RNA4 is by far the most abundant (+) stranded RNA of BMV and related viruses. These observations are not consistent with a role for (−) RNA4 as a template for (+) strand synthesis. Instead, it may be a dead-end product derived from the origin of replication at the 3′ end of the RNA4 (+) strand. In summary, proving a premature termination mechanism will not be straightforward. Most likely use of a cell-free replicase extract will be necessary, to clearly show whether this mechanism is used by viruses such as RCNMV.

Subgenomic promoters can vary greatly between and even within viruses. For example, all three BYDV sgRNA promoters have widely different primary and secondary structures and positions relative to the start site (Koev *et al.,* 1999; Koev and Miller, 2000). The only common feature is the hexanucleotide GUGAAG [in (+) sense] at the 5′ ends of sgRNAs 1 and 2. The sgRNA1 promoter is located mostly upstream of the initiation site. It contains an essential stem–loop and one that hinders sgRNA synthesis (Koev *et al.,* 1999). In contrast, sgRNA2 and sgRNA3 core promoters are located downstream of the sgRNA start sites. Such an arrangement of transcription regulatory elements suggests recognition by different sites in the replicase or by different host factors. The sgRNAs may even be made by different mechanisms.

A potential driving force behind the evolution of divergent promoters within the same RNA virus may be the limited size of viral genomes, which often requires overlapping of genes and *cis*-acting elements. Subgenomic promoters may have been forced to share genomic locations with other important elements or ORFs, which would require mutual compromises. The flexibility in sgRNA promoter structure and perhaps mechanism would maximize opportunities to overlap with other functions. sgRNA promoters could arise within a sequence that already has a function, such as an ORF. Thus, sgRNA promoters may have evolved independently many times.

In all viruses that produce multiple sgRNAs, the 3′-proximal sgRNAs tend to be the most abundant (e.g., Kelly *et al.,* 1994). This effect is especially extreme when the same sgRNA promoter is duplicated artificially (French and Ahlquist, 1988; Koev *et al.,* 1999). We propose that, as the replicase proceeds toward the 5′ end of genomic RNA template during (−) strand synthesis, immediately upon synthesizing a sgRNA promoter, a protein becomes associated with the (−) strand promoter. This serves as a flag to attract other replicases to initiate (+) strand sgRNA synthesis. If the amount of this protein in the cell were limiting, or if free replicases were very quick to recognize the tagged promoter, initiation would occur most frequently at the 5′-proximal [in (−) sense] promoters. Another possibility would be that the promoter-recognition protein is initially associated with the replicase and then left behind on the (−) strand promoter as it is synthesized. This process would not be completely efficient, so that some replicases would not shed this sgRNA promoter recognition factor until the next

promoter is reached, and so on for the next promoter upstream. An alternative explanation for the abundance of 3′-proximal sgRNAs would be that a portion of the replicases turn around and begin synthesizing (+) sgRNA off the nascent (−) strand as soon as the sgRNA promoter is made. However, this would lead to an enrichment for slightly larger-than-sgRNA-length (−) strands, which have not been observed. Also, this would be difficult to reconcile with the vast excess of (+) strand compared to (−) strand sgRNA that is generated. The diverse promoter sequences in multi-sgRNA viruses may have different affinities for the sgRNA promoter recognition factors, which could allow differential regulation of sgRNA synthesis, independent of genome position.

Because sgRNAs of *Nidovirales* contain leader sequences derived from the 5′ ends of gRNAs, they must be synthesized by discontinuous transcription (Lai and Cavanagh, 1997). Two different versions of discontinuous RNA transcription have been suggested for sgRNA synthesis in *Nidovirales* (Fig. 1). Leader-primed transcription proposes that viral replicase produces free plus strand leader RNAs at 90 nt (Lai *et al.,* 1984; Zhang and Lai, 1994). These leader molecules then prime (+) sgRNA synthesis at the short complementary regions in the IG regions of the (−) gRNA that serve as sgRNA "promoters." In MHV these sequences are UCUAA in the leader and UUAGA in the (−) strand IG region. However, UCUAA is not at the 3′ end of the (+) strand leader transcripts (Lai *et al.,* 1984). Thus, a proofreading nuclease would have to remove the mismatched 3′ bases before priming could occur. There is no evidence of such an activity, but it resembles a phenomenon in some retroelements in which reverse transcription is primed by an internal region in a tRNA. The 3′-terminal 36 bases of the tRNA do not anneal to the primer binding site on the retrotransposon Ty5 RNA and must be removed, by an as yet unknown mechanism, prior to primer extension (Ke *et al.,* 1999). If any RNA viruses were to have proofreading activity, coronaviruses would be good candidates. This would explain the ability of coronaviruses to maintain such large genomes, without accumulation of deleterious mutations. Mutagenesis studies have demonstrated that this base pairing is not the only transcription determinant and that adjacent sequence influences transcription (Chang *et al.,* 1996; van Marle *et al.,* 1999a). Furthermore, a MHV DI RNA that lacks IG consensus sequences still produces sgRNAs, but with very heterogeneous leader–sgRNA junctions (Fischer *et al.,* 1997). The mechanism may be protein-mediated (below) with base pairing serving only to make a more precise alignment.

An alternative model is that discontinuous sgRNA transcription occurs during (−) strand RNA synthesis (Sawicki and Sawicki, 1990). The requirement for base pairing between leader and the (−) intergenic region was demonstrated clearly by van Marle *et al.* (1999a) in equine arteritis arterivirus (EAV), which probably uses the same mechanism as coronaviruses. They mutated the IG region and the 5′ leader sequence to show that base pairing is necessary and sufficient for synthesis, independent of sequence. These authors favor the mechanism of replicase hopping in (−) strand synthesis: the replicase remains attached to the nascent (−) strand, which releases from the IG region of the (+) template and re-pairs with the homologous sequence in the leader sequence (which is in an exposed loop at the end of a long stem). (−) strand transcription then proceeds to the 5′ end of the genome (Fig. 1B). This mechanism resembles the copy-choice RNA recombination mechanism and requires no proofreading nuclease activity. This mechanism is also supported by kinetic studies showing that the sgRNAs of MHV accumulated in direct proportion to subgenomic-sized (−) strands and RFs (Baric and Yount, 2000). Lai and Cavanagh (1997) propose that both mechanisms may take place, because MHV sgRNA accumulation has a genomic-length UV target size early in infection, but a subgenomic-length UV target size later in infection (Yokomori *et al.,* 1992).

To understand mechanistic details of sgRNA synthesis in the absence of replication, RNA degradation, and other events in an infected cell, cell-free replicase extracts have been used. The best-characterized subgenomic promoter is that of BMV RNA4 which is transcribed from genomic RNA3. Mapped to a region that includes 74 to 95 nt upstream and 16 nt downstream of the initiation site, it contains several functional elements, a combination of which provides full transcription activity *in vivo* (French and Ahlquist, 1988). A template-dependent, template-specific, cell-free extract capable of subgenomic RNA4 synthesis from (−) strand RNA3 (Miller *et al.,* 1985) was used to map the subgenomic promoter to a 62-nt region (Marsh *et al.,* 1988). Kao's group has dissected this promoter to near atomic resolution, using a more purified RdRp preparation. They found that a 21-nt promoter element is sufficient for the basal level of sgRNA synthesis by BMV RdRp (Adkins *et al.,* 1997). Only the bases at positions −17, −14, −13, and −11 relative to the start site at +1 are required. These are part of a consensus sequence shared by subgenomic promoters of the *Bromoviridae* and alphaviruses: 3′-**AGA**-NN**C**N**G**$(N)_5$**A**$(N)_{3-5}$X**A**-5′, where X, which is usually a U or a C, is the initiation site (Siegel *et al.,* 1997). The cowpea chlorotic mottle bromovirus sgRNA promoter also requires the same four bases as BMV, plus additional bases at −20, −16, −15, and −10 (Adkins and Kao, 1998). In contrast, Haasnoot *et al.* (2000) found that the conserved G at −12 (equivalent to G −11 in BMV) is not required in the alfalfa mosaic virus (AMV, *Bromoviridae*) subgenomic promoter. Moreover, they found that a hairpin loop that incorporates many of the consensus bases is required in both AMV and BMV sgRNA promoters (Haasnoot *et al.,* 2000). The unrelated turnip crinkle

virus (*Tombusviridae*) has two sgRNA promoters. They share 40% overall sequence similarity including short conserved sequences at the transcription start sites and similar hairpins predicted to form immediately upstream of the sgRNA 5′ end (Wang *et al.,* 1999). Flexibility in promoter recognition led Adkins and Kao (1998) to speculate that the RNA induces the replicase to recognize different nucleotide contacts in the promoter. This could explain the enigma of how a replicase recognizes widely different promoters in the same viral genome, including the entirely different tRNA-like structure at the 3′ end of the BMV (+) strand (Dreher and Hall, 1988).

The minimal promoter requirements *in vitro* are generally insufficient *in vivo* (French and Ahlquist, 1988). Additional sequences needed for sgRNA synthesis *in vivo* probably act by allowing the replicase complex to come into proximity with the core promoter or present it in a proper fashion making it accessible to the replicase. Such requirements may not be critical in highly purified *in vitro* systems where core promoter and replicase are more readily accessible to each other.

A fundamental question is how does the replicase recognize the sgRNA promoter, in contrast to the origins of genomic RNA (−) and (+) strand synthesis? Initiation at these three competing, yet structurally dissimilar sites must be precisely regulated. Subgenomic promoters generally do not resemble the genomic origins, so either the replicase has a multifunctional binding site and separate RNA binding sites for each promoter or separate RNA-binding proteins exist for sgRNA promoters. Obviously the virus-encoded RNA-dependent RNA polymerase is required for sgRNA transcription, but other proteins in the replicase holoenzyme, or as separate entities, specifically modulate sgRNA synthesis.

Alphavirus proteins nsp2 and nsp3 affect synthesis of the single 26S subgenomic RNA. nsp1, nsp2, nsp3, and (via leaky stop codon read-through) nsp4 are produced as a single protein which subsequently undergoes proteolytic cleavage into the individual subunits. The uncleaved nsp1–2–3 complex, along with nsp4, is capable of (−) strand synthesis, but nsp2 and nsp3 must exist as separate proteins for synthesis of genomic and subgenomic (+) strands (Lemm *et al.,* 1994). A temperature-sensitive mutation in nsp2 of Semliki Forest virus resulted in greatly reduced 26S sgRNA, reduced subgenomic RFs, and disassociation of nsp2 from the membranous cellular fraction on which replication is believed to occur (Suopanki *et al.,* 1998). Thus, nsp2 has all the hallmarks of a protein that links the replicase to the sgRNA promoter (Suopanki *et al.,* 1998). A mutation in nsp3 reduced sgRNA synthesis by Sindbis virus, even though the authors doubt that it interacts directly with the promoter or affects proteolytic cleavage from the larger form (LaStarza *et al.,* 1994). Complex interactions and conformational changes between the nonstructural proteins are probably important for regulating genomic vs subgenomic RNA synthesis in the alphaviruses (LaStarza *et al.,* 1994).

In EAV arterivirus, a single mutation (Ser2429Pro) in the nsp10 proteolytic fragment of the RdRp ORF drastically reduced (+) and (−) sgRNA synthesis but had no effect on genomic RNA replication (van Marle *et al.,* 1999b). The role of Ser2429 is unknown, but it is not required for normal proteolytic processing or phosphorylation of the protein.

The coat protein of AMV plays a positive role in sgRNA synthesis. Reusken *et al.* (1997) found a correlation between mutations in the 3′ end of AMV RNA3 that blocked coat protein binding and specific loss of sgRNA4 accumulation. Mutations in the small, 3′-terminal stem–loop of the (+) strand reduced sgRNA (+) strand accumulation much more than genomic (+) strand, but had no effect on (−) strand synthesis *in vitro* (Reusken *et al.,* 1997). These mutations also prevent coat protein binding which is essential for infection *in vivo.*

Host proteins may be involved in direct sgRNA promoter recognition, although few have been identified. For a detailed review on host proteins involved in RNA virus replication, see Lai (1998). One candidate host protein for coronavirus sgRNA synthesis is hnRNP A1. It specifically binds the (−) IG sequence and the (−) strand of the conserved genomic 5′ leader sequence (Li *et al.,* 1997). Pyrimidine tract binding protein (PTB) binds the (+) strand of the 5′ leader (Li *et al.,* 1999). hnRNP A1 and PTB interact in mRNA splicing, so it is conceivable that they interact to facilitate leader priming of sgRNA synthesis (Li *et al.,* 1999). PTB also binds the (−) strand of the 3′ UTR (Huang and Lai, 1999). This causes a conformational change in the RNA that correlates with sgRNA transcription activity (Huang and Lai, 1999). Thus, distant regions at both ends of the coronavirus genome appear to regulate sgRNA synthesis.

It may seem counterintuitive that host proteins specifically recognize viral RNA sequences. Why would the host have proteins ready and waiting to support replication of a specific virus? The answer may lie in the fact that both the error rate of genome replication and the generation time of the virus are several orders of magnitude greater than those of the host. Thus, the virus can adapt to the host faster than the host can escape recognition by the virus. Virus infection may be like a SELEX (systematic evolution of ligands by exponential enrichment) reaction in which the selection occurs during virus evolution for viral RNA sequences that bind to and exploit host proteins with maximum efficiency. This was demonstrated by *in vitro* SELEX in which RNAs that bound the replicase of RNA bacteriophage Qβ were selected from a large pool of random RNA sequences (Brown and Gold, 1996). Two populations were obtained which structurally and functionally resembled the replicase binding sites in the viral (+) and (−) strands. Both of these sites were shown to be specifically recognized

by host-encoded subunits of the replicase (Brown and Gold, 1996).

It seems that evolution has solved the "problem" of how to synthesize viral sgRNAs by a remarkable variety of mechanisms and control signals. However, much additional research is necessary to understand the mechanisms. (1) Proof of termination during (−) synthesis for those viruses that may use this mechanism still awaits. (2) *In vitro* systems would help dissection of the mechanism and roles of proteins, but they have been difficult to obtain, probably due to membrane association of most replication complexes. Also, *in vitro* systems may lack important properties found only *in vivo.* (3) Identification of host proteins involved in the process is essential to understanding the mechanisms of subgenomic and genomic RNA replication. Genetic approaches have been hindered by the essential nature (to the host) of many host proteins used by the virus. Approaches such as transferring an entire viral replicon into yeast may provide a tractable host genetic system (Ishikawa *et al.,* 1997). Certainly, enough tools are in-hand to assure exciting future research, as the diverse mechanisms that balance RNA viral genomic replication and subgenomic RNA transcription are revealed.

ACKNOWLEDGMENTS

The authors thank Dr. Michael M. C. Lai for critical reading of the manuscript. This work was funded by the USDA National Research Initiative, Grant 98-35303-6447, and the Iowa Agricultural and Home Economics Experiment Station (Journal Paper J-18851; Project 3545).

REFERENCES

Adkins, S., and Kao, C. C. (1998). Subgenomic RNA promoters dictate the mode of recognition by bromoviral RNA-dependent RNA polymerases. *Virology* **252,** 1–8.

Adkins, S., Siegel, R. W., Sun, J. H., and Kao, C. C. (1997). Minimal templates directing accurate initiation of subgenomic RNA synthesis *in vitro* by the brome mosaic virus RNA-dependent RNA polymerase. *RNA* **3,** 634–647.

Balmori, E., Gilmer, D., Richards, K., Guilley, H., and Jonard, G. (1993). Mapping the promoter for subgenomic RNA synthesis on beet necrotic yellow vein virus RNA 3. *Biochimie* **75,** 517–521.

Baric, R. S., and Yount, B. (2000). Subgenomic negative-strand RNA function during mouse hepatitis virus infection. *J. Virol.* **74,** 4039–4046.

Barr, J. N., Whelan, S. P., and Wertz, G. W. (1997) *cis*-Acting signals involved in termination of vesicular stomatitis virus mRNA synthesis include the conserved AUAC and the U7 signal for polyadenylation. *J. Virol.* **71,** 8718–8725.

Baulcombe, D. C. (1999). Fast forward genetics based on virus-induced gene silencing. *Curr. Opin. Plant Biol.* **2,** 109–113.

Boccard, F., and Baulcombe, D. (1993). Mutational analysis of *cis*-acting sequences and gene function in RNA3 of cucumber mosaic virus. *Virology* **193,** 563–578.

Brown, D., and Gold, L. (1996). RNA replication by Qβ replicase: A working model. *Proc. Natl. Acad. Sci. USA* **93,** 11558–11562.

Chang, R. Y., Krishnan, R., and Brian, D. A. (1996). The UCUAAAC promoter motif is not required for high-frequency leader recombination in bovine coronavirus defective interfering RNA. *J. Virol.* **70,** 2720–2729.

Dreher, T. W., and Hall, T. C. (1988). Mutational analysis of the sequence and structural requirements in brome mosaic virus RNA for (−) strand promoter activity. *J. Mol. Biol.* **201,** 31–40.

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

French, R., and Ahlquist, P. (1988). Characterization and engineering of sequences controlling *in vivo* synthesis of brome mosaic virus subgenomic RNA. *J. Virol.* **62,** 2411–2420.

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

Hahn, C. S., Hahn, Y. S., Braciale, T. J., and Rice, C. M. (1992). Infectious Sindbis virus transient expression vectors for studying antigen processing and presentation. *Proc. Natl. Acad. Sci. USA* **89,** 2679–2683.

Huang, P., and Lai, M. M. C. (1999). Polypyrimidine tract-binding protein binds to the complementary strand of the mouse hepatitis virus 3′ untranslated region, thereby altering RNA conformation. *J. Virol.* **73,** 9110–9116.

Ishikawa, M., Janda, M., Krol, M. A., and Ahlquist, P. (1997). *In vivo* DNA expression of functional brome mosaic virus RNA replicons in *Saccharomyces cerevisiae. J. Virol.* **71,** 7781–7790.

Jaspars, E. M. (1998). A core promoter hairpin is essential for subgenomic RNA synthesis in alfalfa mosaic alfamovirus and is conserved in other Bromoviridae. *Virus Genes* **17,** 233–242.

Johnson, B. W., Olson, K. E., Allen-Miura, T., Rayms-Keller, A., Carlson, J. O., Coates, C. J., Jasinskiene, N., James, A. A., Beaty, B. J., and Higgs, S. (1999). Inhibition of luciferase expression in transgenic *Aedes aegypti* mosquitoes by Sindbis virus expression of antisense luciferase RNA. *Proc. Natl. Acad. Sci. USA* **96,** 13399–13403.

Johnston, J. C., and Rochon, D. M. (1995). Deletion analysis of the promoter for the cucumber necrosis virus 0.9-kb subgenomic RNA. *Virology* **214,** 100–109.

Ke, N., Gao, X., Keeney, J. B., Boeke, J. D., and Voytas, D. F. (1999). The yeast retrotransposon Ty5 uses the anticodon stem–loop of the initiator methionine tRNA as a primer for reverse transcription. *RNA* **5,** 929–938.

Kelly, L., Gerlach, W. L., and Waterhouse, P. M. (1994). Characterisation of the subgenomic RNAs of an Australian isolate of barley yellow dwarf luteovirus. *Virology* **202,** 565–573.

Kim, K.-H., and Hemenway, C. (1996). The 5′ nontranslated region of potato virus X RNA affects both genomic and subgenomic RNA synthesis. *J. Virol.* **70,** 5533–5540.

Kim, K. H., and Hemenway, C. L. (1999). Long-distance RNA–RNA interactions and conserved sequence elements affect potato virus X plus-strand RNA accumulation. *RNA* **5,** 636–645.

Koev, G., and Miller, W. A. (2000). A positive strand RNA virus with three very different subgenomic RNA promoters. *J. Virol.* **74,** in press.

Koev, G., Mohan, B. R., and Miller, W. A. (1999). Primary and secondary structural elements required for synthesis of barley yellow dwarf virus subgenomic RNA1. *J. Virol.* **73,** 2876–2885.

Koonin, E. V., and Dolja, V. V. (1993). Evolution and taxonomy of positive-strand RNA viruses: Implications of comparative analysis of amino acid sequences. *Crit. Rev. Biochem. Mol. Biol.* **28,** 375–430.

Kroner, P. A., Young, B. M., and Ahlquist, P. (1990). Analysis of the role of brome mosaic virus 1a protein domains in RNA replication, using linker insertion mutagenesis. *J. Virol.* **64,** 6110–6120.

Kumagai, M. H., Donson, J., della-Cioppa, G., Harvey, D., Hanley, K., and Grill, L. K. (1995). Cytoplasmic inhibition of carotenoid biosynthesis with virus-derived RNA. *Proc. Natl. Acad. Sci. USA* **92,** 1679–1683.

Lai, M. M., Baric, R. S., Brayton, P. R., and Stohlman, S. A. (1984). Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus. *Proc. Natl. Acad. Sci. USA* **81,** 3626–3630.

Lai, M. M. C. (1998). Cellular factors in the transcription and replication

of viral RNA genomes: A parallel to DNA-dependent RNA transcription. *Virology* **244,** 1–12.

Lai, M. M. C., and Cavanagh, D. (1997). The molecular biology of coronaviruses. *Adv. Virus Res.* **48,** 1–100.

LaStarza, M. W., Lemm, J. A., and Rice, C. M. (1994). Genetic analysis of the nsP3 region of Sindbis virus: Evidence for roles in minus-strand and subgenomic RNA synthesis. *J. Virol.* **68,** 5781–5791.

Lehto, K., Grantham, G. L., and Dawson, W. O. (1990). Insertion of sequences containing the coat protein subgenomic RNA promoter and leader in front of the tobacco mosaic virus 30K ORF delays its expression and causes defective cell-to-cell movement. *Virology* **174,** 145–157.

Lemm, J. A., Rumenapf, T., Strauss, E. G., Strauss, J. H., and Rice, C. M. (1994). Polypeptide requirements for assembly of functional Sindbis virus replication complexes: A model for the temporal regulation of minus- and plus-strand RNA synthesis. *EMBO J.* **13,** 2925–2934.

Levis, R., Schlesinger, S., and Huang, H. V. (1990). Promoter for Sindbis virus RNA-dependent subgenomic RNA transcription. *J. Virol.* **64,** 1726–1733.

Li, H. P., Huang, P., Park, S., and Lai, M. M. (1999). Polypyrimidine tract-binding protein binds to the leader RNA of mouse hepatitis virus and serves as a regulator of viral transcription. *J. Virol.* **73,** 772–777.

Li, H. P., Zhang, X., Duncan, R., Comai, L., and Lai, M. M. (1997). Heterogeneous nuclear ribonucleoprotein A1 binds to the transcription-regulatory region of mouse hepatitis virus RNA. *Proc. Natl. Acad. Sci. USA* **94,** 9544–9549.

Lin, Y.-J., Zhang, X., Wu, R.-C., and Lai, M. M. C. (1996). The 3′ untranslated region of coronavirus RNA is required for subgenomic mRNA transcription from a defective interfering RNA. *J. Virol.* **70,** 7236–7240.

Marsh, L. E., Dreher, T. W., and Hall, T. C. (1988). Mutational analysis of the core and modulator sequences of the BMV RNA3 subgenomic promoter. *Nucleic Acids Res.* **16,** 981–995.

McCormick, A. A., Kumagai, M. H., Hanley, K., Turpen, T. H., Hakim, I., Grill, L. K., Tuse, D., Levy, S., and Levy, R. (1999). Rapid production of specific vaccines for lymphoma by expression of the tumor-derived single-chain Fv epitopes in tobacco plants. *Proc. Natl. Acad. Sci. USA* **96,** 703–708.

Meulewaeter, F., Cornelissen, M., and van Emmelo, J. (1992). Subgenomic RNAs mediate expression of cistrons located internally in the genomic RNA of tobacco necrosis virus strain A. *J. Virol.* **66,** 6419–6428.

Miller, W. A., Dreher, T. W., and Hall, T. C. (1985). Synthesis of brome mosaic virus subgenomic RNA *in vitro* by internal initiation on (−) sense genomic RNA. *Nature* **313,** 68–70.

Miller, W. A., Brown, C. M., and Wang, S. (1997). New punctuation for the genetic code: Luteovirus gene expression. *Semin. Virol.* **8,** 3–13.

Nagy, P. D., Zhang, C., and Simon, A. E. (1998). Dissecting RNA recombination *in vitro:* Role of RNA sequences and the viral replicase. *EMBO J.* **17,** 2392–2403.

Reusken, C. B. E. M., Neeleman, L., Brederode, F. T., and Bol, J. F. (1997). Mutations in coat protein binding sites of alfalfa mosaic virus RNA 3 affect subgenomic RNA 4 accumulation and encapsidation of viral RNAs. *J. Virol.* **71,** 8385–8391.

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

Shivprasad, S., Pogue, G. P., Lewandowski, D. J., Hidalgo, J., Donson, J., Grill, L. K., and Dawson, W. O. (1999). Heterologous sequences greatly affect foreign gene expression in tobacco mosaic virus-based vectors. *Virology* **255,** 312–323.

Siegel, R. W., Adkins, S., and Kao, C. C. (1997). Sequence-specific recognition of a subgenomic RNA promoter by a viral RNA polymerase. *Proc. Natl. Acad. Sci. USA* **94,** 11238–11243.

Sit, T. L., Vaewhongs, A. A., and Lommel, S. A. (1998). RNA-mediated transactivation of transcription from a viral RNA. *Science* **281,** 829–832.

Suopanki, J., Sawicki, D. L., Sawicki, S. G., and Kaariainen, L. (1998). Regulation of alphavirus 26S mRNA transcription by replicase component nsP2. *J. Gen. Virol.* **79,** 309–319.

van der Kuyl, A. C., Langereis, K., Houwing, C. J., Jaspars, E. M., and Bol, J. F. (1990). *cis*-Acting elements involved in replication of alfalfa mosaic virus RNAs *in vitro*. *Virology* **253,** 327–336.

van der Vossen, E. A., Notenboom, T., and Bol, J. F. (1995). Characterization of sequences controlling the synthesis of alfalfa mosaic virus subgenomic RNA *in vivo*. *Virology* **212,** 663–672.

van Dinten, L. C., den Boon, J. A., Wassenaar, A. L., Spaan, W. J., and Snijder, E. J. (1997). An infectious arterivirus cDNA clone: Identification of a replicase point mutation that abolishes discontinuous mRNA transcription. *Proc. Natl. Acad. Sci. USA* **94,** 991–996.

van Marle, G., Luytjes, W., van der Most, R. G., van der Straaten, T., and Spaan, W. J. (1995). Regulation of coronavirus mRNA transcription. *J. Virol.* **69,** 7851–7856.

van Marle, G., Dobbe, J. C., Gultyaev, A. P., Luytjes, W., Spaan, W. J., and Snijder, E. J. (1999a). Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. *Proc. Natl. Acad. Sci. USA* **96,** 12056–12061.

van Marle, G., van Dinten, L. C., Spaan, W. J., Luytjes, W., and Snijder, E. J. (1999b). Characterization of an equine arteritis virus replicase mutant defective in subgenomic mRNA synthesis. *J. Virol.* **73,** 5274–5281.

Wang, J., Carpenter, C. D., and Simon, A. E. (1999). Minimal sequence and structural requirements of a subgenomic RNA promoter for turnip crinkle virus. *Virology* **253,** 327–336.

Wang, J., and Simon, A. E. (1997). Analysis of the two subgenomic RNA promoters for turnip crinkle virus *in vivo* and *in vitro*. *Virology* **232,** 174–186.

Yokomori, K., Banner, L. R., and Lai, M. M. C. (1992). Coronavirus mRNA transcription: UV light transcriptional mapping studies suggest an early requirement for a genomic-length template. *J. Virol.* **66,** 4671–4678.

Zavriev, S. K., Hickey, C. M., and Lommel, S. A. (1996). Mapping of the red clover necrotic mosaic virus subgenomic RNA. *Virology* **216,** 407–410.

Zhang, G., Slowinski, V., and White, K. A. (1999). Subgenomic mRNA regulation by a distal RNA element in a (+)-strand RNA virus. *RNA* **5,** 550–561.

Zhang, X., and Lai, M. M. (1994). Unusual heterogeneity of leader–mRNA fusion in a murine coronavirus: Implications for the mechanism of RNA transcription and recombination. *J. Virol.* **68,** 6626–6633.