How a Small DNA Virus Uses dsRNA but Not RNAi to Regulate Its Life Cycle

R. Gu, Z. Zhang,* and G.G. Carmichael

Department of Genetics and Developmental Biology, University of Connecticut Health Center, Farmington, Connecticut 06030

Mouse polyomavirus contains a circular DNA genome, with early and late genes transcribed from opposite strands. At early times after infection, genes encoded from the early transcription unit are predominantly expressed. After the onset of viral DNA replication, expression of genes encoded from the late transcription unit increases dramatically. At late times, late primary transcripts are inefficiently polyadenylated, leading to the generation of multigenomic RNAs that are precursors to mature mRNAs. These transcripts contain sequences complementary to the early RNAs and down-regulate early-strand gene expression by inducing RNA editing. Our recent work leads to a model where the production of the multigenomic late RNAs is also controlled by the editing of poly(A) signals, directed by overlapping primary transcripts.

Most DNA viruses exhibit striking temporal regulation of the expression of their genes. Viral genes involved in altering host functions or in regulating viral replication are expressed at early times, whereas virion structural proteins are only expressed at high levels after the onset of viral DNA replication. For some of these viruses (e.g., human adenoviruses and Simian Virus 40) (see Lewis and Manley 1985; Wiley et al. 1993), the temporal regulation of gene expression appears to be at the level of transcription initiation. Work in our laboratory has suggested that an interesting and fundamentally different type of regulation has an important role in the early-to-late switch in gene expression in cells infected with the murine polyomavirus.

Polyomavirus lytically infects mouse cells in tissue culture and is an apparently harmless passenger virus in wild mouse populations. In other rodents, it is tumorigenic and efficiently transforms rat or hamster cells in culture (Tooze 1980). Due to its small genome size and ease of manipulation, it provides us with a good model system for studying not only the molecular biology of cell transformation and tumorigenesis, but also the mechanisms of regulation of eukaryotic gene expression. We have been interested in understanding how viral RNA molecules are made and processed in infected cells. Late viral gene expression has a number of unusual features that have turned out to be useful for helping us to unravel fundamental aspects of RNA synthesis, processing, regulation, and mRNA transport from the nucleus.

**GENOME ORGANIZATION**

The polyoma genome is a circular DNA molecule of about 5300 bp. Our laboratory strain, 59RA, is 5327 bp (Ruley and Fried 1983). The genome is divided into “early” and “late” regions, which are expressed and regulated differently as infection proceeds (Fig. 1A) (Griffin and Fried 1975; Kamen et al. 1975, 1980a, b). The early and late transcription units extend in opposite directions around the circular genome from start sites near the unique, bidirectional origin of DNA replication (Crawford et al. 1974; Griffin and Fried 1975). Primary RNA products from the early transcription unit are alternatively spliced to yield three early mRNAs that code for the large T antigen (100 kD), the middle T antigen (56 kD), and the small T antigen (22 kD). Large T binds to sequences in or near the DNA replication origin region (Gaudray et al. 1981; Pomerantz et al. 1983; Cowie and Kamen 1984; Dilworth et al. 1984) and is involved in the initiation of DNA replication, indirectly in the autoregulation of early-strand RNA levels (Cogen 1978; Farmerie and Folk 1984; Liu and Carmichael 1993), and indirectly in the activation of high levels of expression from the late promoter (Cahill et al. 1990; Liu and Carmichael 1993). The other two early proteins are dispensable for lytic infection, but they are important for cell transformation. Late primary transcripts accumulate after the onset of DNA replication and are also spliced in alternative ways to give mRNAs that code for the three viral structural proteins VP1, VP2, and VP3.

**TEMPORAL REGULATION OF GENE EXPRESSION**

Gene expression during lytic infection of permissive mouse cells proceeds in a well-defined temporally regulated manner (Kamen et al. 1975; Beard et al. 1976; Piper 1979). Immediately after infection, RNA from the early transcription unit (E-RNA) begins to accumulate; however, RNA from the late transcription unit (L-RNA) accumulates more slowly. At 12–15 hours after infection, the early–late RNA ratio is about 4 to 1 (see Fig. 1B) (Kamen et al. 1975; Piper 1979; Hyde-DeRuyoucher and Carmichael 1988) and in the presence...
of DNA replication inhibitors, the ratio is 10 to 1 or higher. At 12–15 hours postinfection, viral DNA replication commences and L-RNA begins to accumulate rapidly, whereas E-RNA accumulates at a slower rate. Thus, there is a dramatic change in the relative abundances of E-RNA and L-RNA; by 24 hours postinfection, the early-to-late RNA ratio is as low as 1 to 50 (Kamen et al. 1975; Piper 1979; Hyde-DeRuyscher and Carmichael 1988). This early–late “switch” is dependent on viral DNA replication; if replication is inhibited, E-RNA accumulates to abnormally high levels with minimal accumulation of L-RNA (Cogen 1978; Heiser and Eckhart 1982; Kamen et al. 1982; Farmerie and Folk 1984; Hyde-DeRuyscher and Carmichael 1988). Figure 1B illustrates the early–late switch using an RNase protection assay, and Figure 1C presents an idealized depiction of it.

We have generated data that have uncovered unexpected mechanisms about how late genes are activated after the onset of viral DNA replication, as well as how early genes are down-regulated at late times. Furthermore, this work has led to new insights into the mechanism of action of antisense RNA in cells. It has been commonly accepted in the field for many years that the early–late switch is the result of T antigen repression of the early promoter, coupled with a trans-activation of the late promoter. In fact, there is little experimental support for this notion, and it is not true. We have shown that this temporally regulated switch is not controlled mainly at the level of transcription initiation, but results from changes in transcription elongation and/or RNA stability (Hyde-DeRuyscher and Carmichael 1988, 1990; Liu and Carmichael 1993; Liu et al. 1994). At early times, late-strand transcripts are produced. However, these are inefficiently spliced and exported to the cytoplasm, and they have a short half-life in the nucleus (Fig. 2A) (Hyde-DeRuyscher and Carmichael 1988, 1990; Liu and Carmichael 1993; Liu et al. 1994). At late times, late nuclear RNAs are heterogeneous in size, the result of inefficient polyadenylation and transcription termination, and range from about 2.5 kb to more than 60 kb in length (Fig. 2B) (Acheson et al. 1971; Acheson 1976, 1978; Birg et al. 1977; Treisman and Kamen 1981). Most late RNA sequences never leave the nucleus as they are removed during mRNA processing and are subsequently degraded (Acheson 1976, 1984).
Late-strand pre-mRNA molecules are processed into mature mRNAs using a highly unusual pathway that involves inefficient polyadenylation and ordered splice site selection from precursors containing tandemly repeated introns and exons. Unlike early primary transcripts, late messages contain at their 5′ ends multiple tandem repeats of the 57-base noncoding late leader exon, which appears only once in the viral genome. Pre-mRNA molecules are processed by a pathway that includes the splicing of late leader exons to each other (Fig. 3A). We have shown that each class of late viral message (encoding virion structural proteins VP1, VP2, or VP3) consists of molecules with between 1 and 12 tandem leader units at their 5′ ends (Hyde-DeRuyscher and Carmichael 1990).

The life cycle of the virus appears to be connected to the processing of late pre-mRNAs, and this processing is in turn related to the inefficient use of the late polyadenylation signal. This can be seen in Figure 3B. Note that the great majority of late messages contain multiple tandem leaders at their 5′ ends. This reflects the inefficiency of late polyadenylation, which has been estimated to be about 50% each time RNA polymerase II traverses the poly(A) signal (Hyde-DeRuyscher and Carmichael 1990; Liu and Carmichael 1993). Second, dsRNA formation in the nucleus would be expected to lead to editing by the ADAR enzyme. In the nucleus, adenosine residues in dsRNAs can be edited by ADAR to inosines by a process of hydrolytic deamination (Bass 2002). We have shown that this indeed happens in a polyomavirus infection. At late times, early-strand RNAs are extensively edited, with many transcripts exhibiting about 50% of their adenosines converted to inosines. Furthermore, these promiscuously edited RNAs are quantitatively retained in the nucleus and therefore are not translated into mutant proteins in the cytoplasm (Kumar and Carmichael 1997). Thus, RNA editing dramatically reduces the amount of cytoplasmic translatable early-strand mRNAs at late times. On the other hand, edited late-strand sequences lie within the large intron that is removed and degraded in the nucleus, so that editing does not directly affect late-strand mRNAs.

**WHAT IS THE SWITCH THAT ACTIVATES LATE-STRAND SPlicing?**

The polyoma early–late switch is clearly connected to the inefficient use of the late poly(A) signal. But what is the basis for this inefficiency? We have carried out many experiments to uncover the cis-acting sequence(s) in the polyoma late region that confers the temporal regulation of gene expression (Barrett et al. 1991; Liu and Carmichael 1993; Batt and Carmichael 1995; Huang and Carmichael 1996). Curiously, experiments aiming to uncover the “culprit” sequences that regulate the change in late polyadenylation efficiency between early and late times in infection have until recently proved frustrating. The results are summarized in Figure 4. There actually

**ANTISENSE REGULATION**

An important contribution to the down-regulation of early RNA levels at late times in infection comes from the multigenomic late-strand transcripts. These RNAs are antisense to the early transcripts and may anneal with them in the nucleus, forming double-stranded RNA (dsRNA). We have developed several lines of evidence pointing to the importance of antisense regulation to the polyoma life cycle. First, mutants that express lower levels of giant late-strand RNAs always exhibit reciprocally increased early RNA levels (Adami et al. 1989; Liu and Carmichael 1993). Second, dsRNA formation in the nucleus would be expected to lead to editing by the ADAR enzyme. In the nucleus, adenosine residues in dsRNAs can be edited by ADAR to inosines by a process of hydrolytic deamination (Bass 2002). We have shown that this indeed happens in a polyomavirus infection. At late times, early-strand RNAs are extensively edited, with many transcripts exhibiting about 50% of their adenosines converted to inosines. Furthermore, these promiscuously edited RNAs are quantitatively retained in the nucleus and therefore are not translated into mutant proteins in the cytoplasm (Kumar and Carmichael 1997).
appear to be no clear sequences within the late region of the virus that are essential for the early–late switch. In all of our studies, only two regions could be deleted without dramatic consequences on the switch. The first is the late leader region. Although the leader exon sequence itself is not important, its flanking splice sites must be functional (Adami et al. 1989). This most likely represents the need for leader-to-leader splicing to allow late RNA accumulation at late times.

The second region important for the early–late switch appears to include the poly(A) signal, which cannot be deleted. However, the late poly(A) region also includes elements essential for early-strand gene expression and thus indirectly for viral DNA replication. We next examined this element in greater detail. First, we asked whether the viral poly(A) signal itself is regulated. Our approach was to determine whether it can be replaced by another, unrelated element. We thus chose to mutationally inactivate the late polyadenylation signal while at the same time inserting a functional “synthetic” poly(A) signal (which obeys the known rules for the composition of an efficient processing signal; see Levitt et al. 1989) that contains some unique restriction sites. The resulting virus is called YZ-3 (Fig. 5). Note that this virus has wild-type open reading frames (ORFs) and differs from wild type only in the late poly(A) region. The early poly(A) signal is completely intact. The mutagenesis involved in the construction of YZ-3 was difficult, as the early and late viral poly(A) signals overlap and the ORFs for viral proteins extend to almost the very end of the mRNAs (Fig. 5A). Interestingly, mutant YZ-3 is viable and grows with kinetics indistinguishable from those of our wild-type virus (Z. Liu and G. Carmichael, unpubl.). This mutant exhibits a normal early–late switch when assayed by either RNase protection assays (Fig. 5B) or reverse transcriptase–polymerase chain reaction (RT-PCR) for leader-to-leader splicing (Fig. 5C). The conclusion from these results is that the normal viral late polyadenylation signal appears to be dispensable for viral growth and for the temporal regulation of early and late gene expression. This led to the conclusion that there appears to be no identifiable element in the viral late region that contributes essentially to the early–late switch.

**A NEW MODEL FOR THE EARLY–LATE SWITCH**

The above results appeared to rule out the late poly(A) signal itself in the regulation of the early–late switch. However, in this region, we noticed an interesting and previously unappreciated feature of the organization of the viral genome. The early and late polyadenylation signals actually overlap, with the primary transcripts potentially overlapping by at least 45 bp (Fig. 6A). This is the case...
for both wild type and YZ-3. Could this overlap be significant? Could overlapping poly(A) signals be regulated by editing? Given what we had already learned about the editing that occurs in the late phase, and that adenosines which are preferred editing targets (those preceded by A’s or U’s) (Polson and Bass 1994; Kumar and Carmichael 1997) are very rich in this overlap region, it seemed reasonable to hypothesize that early and late strands which overlap at their 3′ ends might serve as substrates for editing that is targeted to the polyadenylation signals. In fact, owing to the known nearest-neighbor preferences for ADAR1 editing, the sequence AAUAAA signals. In fact, owing to the known nearest-neighbor preferences for editing (Polson and Bass 1994), and seven edited sequences surrounding the polyadenylation signal (Norbury and Fried 1987) are shown. Note that the sequence AAUAAA is an exceptionally favorable target for ADAR, owing to the well-known nearest-neighbor preferences for editing (Polson and Bass 1994).

What would be the consequence of poly(A)-site editing for the viral life cycle? Interestingly, such editing might be completely consistent with all known results reported so far; a model for the regulation of the polyoma early–late switch by poly(A)-site editing is presented in Figure 6B. If the late poly(A) signal were to become edited at late times, very little transcript overlap occurs, allowing early mRNAs to preferentially accumulate. At late times, overlapping transcripts become edited in the polyadenylation region, leading to a failure of the 3′-processing machinery to recognize the signals. If the late poly(A) signal is edited, transcription proceeds around the genome, with leader–leader splicing removing the large edited intron and resulting in accumulation of late mRNAs. Productive polyadenylation is thus in competition with editing. At the same time, the early poly(A) signal may also become edited. In this case, however, there is no splicing event that can remove the edited sequences, so editing leads to down-regulation of gene expression.

Figure 6. (A) The polyoma early and late polyadenylation signals overlap, with the potential for early-strand and late-strand transcripts to anneal with one another over a region of at least 45 nucleotides. (Dotted lines) Primary transcripts that most likely extend the region of complementarity by an even greater distance. (B) A new model for the early–late switch, in which poly(A)-site editing has a key role. At early times in infection, very little transcript overlap occurs, allowing early mRNAs to preferentially accumulate. At late times, overlapping transcripts become edited in the polyadenylation region, leading to a failure of the 3′-processing machinery to recognize the signals. If the late poly(A) signal is edited, transcription proceeds around the genome, with leader–leader splicing removing the large edited intron and resulting in accumulation of late mRNAs. Productive polyadenylation is thus in competition with editing. At the same time, the early poly(A) signal may also become edited. In this case, however, there is no splicing event that can remove the edited sequences, so editing leads to down-regulation of gene expression.

Figure 7. Poly(A) signals are efficiently edited. RT-PCR was used to amplify edited polyoma sequences as described previously (Kumar and Carmichael 1997), and seven edited sequences surrounding the polyadenylation signal (Norbury and Fried 1987) are shown. Note that the sequence AAUAAA is an exceptionally favorable target for ADAR, owing to the well-known nearest-neighbor preferences for editing (Polson and Bass 1994).
infection. In addition work to be presented elsewhere, we have carried out a number of additional experiments using other approaches to confirm the importance of poly(A)-site overlap for the viral life cycle.

CONCLUSIONS

Taken together, the experiments described above lead us to hypothesize that poly(A)-site overlap and editing is a key regulator of the polyoma early–late switch. If true, this surprising finding would identify poly(A)-site editing as a new form of gene regulation, one that has not been reported previously. Since there exist many nuclear transcripts whose 3′ ends overlap those of nearby genes expressed from the opposite DNA strand, it will be of interest to determine whether poly(A)-site editing might also contribute to the regulation of cellular gene expression.

ACKNOWLEDGMENTS

We are grateful to all members of the laboratory, past and present, who have contributed insights and comments throughout this work. In particular, we acknowledge the contributions of D. Batt, Z. Liu, R. Hyde-DeRuyscher, M. Kumar, and Y. Zhang, and the technical assistance of K. Morris. This work was supported by grants GM066816 and CA045382 from the National Institutes of Health.

REFERENCES


How a Small DNA Virus Uses dsRNA but Not RNAi to Regulate Its Life Cycle

R. GU, Z. ZHANG and G.G. CARMICHAEL

Cold Spring Harb Symp Quant Biol 2006 71: 293-299
Access the most recent version at doi:10.1101/sqb.2006.71.017

References
This article cites 41 articles, 23 of which can be accessed free at:
http://symposium.cshlp.org/content/71/293.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.