RNA (1996), 2:1033–1045. Cambridge University Press. Printed in the USA.

Paramyxovirus RNA editing and the requirement for hexamer genome length

STÉPHANE HAUSMANN,* JEAN-PHILIPPE JACQUES,* and DANIEL KOLAKOFSKY
Department of Genetics and Microbiology, University of Geneva School of Medicine, Geneva, Switzerland

ABSTRACT

Paramyxoviruses cotranscriptionally edit their P gene mRNA by the programmed insertion of G residues into a short G run contained within a larger purine run, via pseudo-templated transcription. The templates for paramyxovirus transcription are genome nucleocapsids in which each nucleoprotein subunit is associated with 6 nt, and only genomes whose lengths are multiples of 6 are found naturally or are replicated efficiently in transfected cell systems. We have examined the effect of varying total genome length on the frequency and number of insertions into the mRNA editing site in a transfected cell system, using constructs that generate mini-genome analogues. We found that, as long as the purine run sequence and the region immediately upstream were unaltered, editing occurred during mRNA synthesis independent of the precise length of the mini-genome. However, when mini-genome constructs whose lengths were not multiples of 6 were used, insertions (or deletions) occurred during antigenome synthesis within the purine run, which strikingly restored the hexamer length. Genome length correction due to changes in the antigenome purine run length occurred only when the mini-genome was not a multiple of 6, and these changes were only poorly affected by mutations in the mRNA editing site and the region immediately upstream. Our results suggest that the mRNA editing site is a natural hotspot for viral polymerase slippage during genome replication, and that this site serves the dual and complementary function of maintaining hexamer genome length. The unusual requirement of paramyxoviruses for genomes of precise hexamer length may have evolved to maintain genome stability against insertions in the mRNA editing site during replication.

Keywords: paramyxovirus; replication; RNA editing; pseudo-templated transcription

INTRODUCTION

Paramyxoviruses contain nonsegmented negative-strand RNA genomes with 6–10 genes, which serve as templates both for monocistronic mRNAs and for full-length antigenomes, the intermediates in genome replication. The paramyxovirus P genes are unusual, in that they generally contain alternate open reading frames (ORFs) that overlap the N-terminus as well as the middle region of the P protein ORF, and express several proteins (Lamb & Kolakofsky, 1996). For Sendai virus (SeV), e.g., the C protein ORF overlaps the N-terminal region of the P ORF and is accessed via ribosomal choice during translational initiation (Curran & Kolakofsky, 1990). The highly conserved, cysteine-rich V ORF, which overlaps the middle of the P ORF, on the other hand, is accessed by a mechanism involving transcriptional choice, which is referred to as cotranscriptional mRNA editing.

The *Paramyxovirinae* are currently organized in three genera, the SeV group (including bovine parainfluenza virus type 3, bPIV3); morbilliviruses (e.g., measles and the distemper viruses); and rubulaviruses (e.g., mumps and SV5). Most of these viral P genes contain an A_nG_n purine run at the start of the internal, overlapping V ORF. mRNAs with expanded G runs are transcribed from these genes in addition to those that are faithful copies of their templates, and the number of G insertions that occur for each virus group mirrors their requirements to switch between the in-frame and out-of-frame ORFs (reviewed in Kolakofsky et al., 1993). For the morbilliviruses and some of the SeV group, which require a +1 frameshift to access the

Reprint requests to: Daniel Kolakofsky, Department of Genetics and Microbiology, University of Geneva School of Medicine, CMU, 9 Ave de Champel, CH1211 Geneva, Switzerland; e-mail: kolakofs@cmu.unige.ch.

* The first two authors contributed equally to the work.

V ORF from the genome-encoded P ORF, a single G is added as the predominant insertional event (left side, Fig. 1). For the rubulaviruses, which require a +2 frameshift to access the remainder of the P ORF from the genone-encoded V ORF, 2 Gs are added at high frequency when insertions occur (right side, Fig. 1). For bPIV3, where both V and another ORF (called D) overlap the middle of the genome-encoded P ORF, 1–6 Gs are added at roughly equal frequency, so that mRNAs encoding all three overlapping ORFs are expressed. Although the available evidence suggests that these Gs are added cotranscriptionally, the term "mRNA editing" has nevertheless been retained to describe these events. RNA editing is defined here as a process in which nucleotide insertion, deletion, or base substitution produces an RNA whose sequence (and informational capacity) differs from that of its template, other than by splicing, 5′ and 3′ end formation, and the creation of hypermodified bases. Both post- and cotranscriptional events are thus included within this definition (reviewed in Cattaneo, 1991; Benne, 1993).

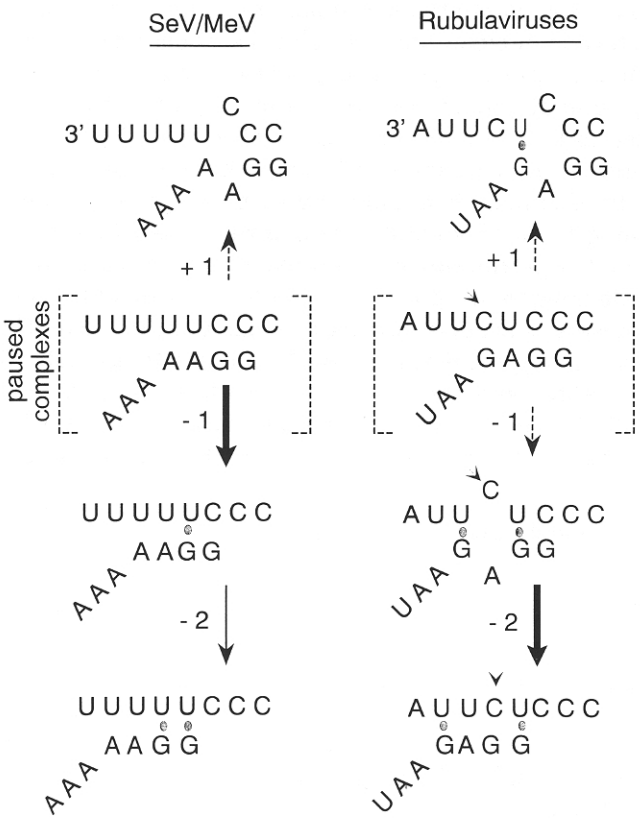

FIGURE 1. Realignment possibilities at the paramyxovirus mRNA editing sites. RNA synthesis complexes of the template (top strand, written 3′ to 5′) and nascent chain (bottom strand, 5′ to 3′) containing four base pairs are shown, for the SeV group and morbilliviruses (left side) and rubulaviruses (right side). The polymerase (not shown) whose catalytic site contains the 3′ end of the nascent (bottom) chain is proposed to pause after incorporation opposite the middle template C residue (second level, in dotted brackets), and the nascent chain to realign on the template, allowing for U:G pairs (highlighted with a shaded circle in between), but not A:C pairs (which are shown looping out). The frequencies of the realignments that occur during natural virus infections are indicated by the strengths of the arrows. Numbers indicate the realignment of the nascent chain; minus as upstream and plus as downstream, i.e., the opposite of the subsequent insertions. For rubulaviruses, the −1 realignment is prevented by the marked (arrowhead) template C residue (which is a U in the SeV/MeV group). Once bypassed by a −2 shift (bottom panel), this template C aligns only with Gs in further shifts.

The G runs of the P gene mRNAs are thought to be expanded by pseudo-templated transcription (Jacques & Kolakofsky, 1991) of the genome C runs as follows: (1) the viral polymerase is postulated to pause before the end of the template C run; (2) the nascent chain, which is base paired to the template by only a few nucleotides, slips backward by one (SeV group and morbilliviruses) or two (rubulaviruses) positions (Fig. 1); and (3) as a result, one or two of the template C residues are thus copied a second time when transcription resumes processively. In the realignment of nascent mRNA and template, U:G (but not A:C) base pairs are permitted and, in analogy to ribosomal frameshifting, the region where alternate base pairing occurs after realignment is called the "slippery sequence" (Jacks & Varmus, 1985; Brierley et al., 1989; Weiss et al., 1990). For most viruses, this cycle of slippage and pseudotemplated synthesis occurs only once, but for bPIV3, it is postulated to occur repeatedly, generating a range of multiple G insertions. The presence of a counting mechanism has therefore been invoked to explain the precise number of G insertions (Pelet et al., 1991).

We have recently described a transfected cell system to study mRNA editing, in which SeV [−] minigenomes (ca. 400 nt in length [cf. Fig. 8], as opposed to the 15.4-kb nondefective SeV genome) are first transcribed from synthetic cDNA constructs by T7 RNA polymerase (provided by infection with a recombinant vaccinia virus, vTF7-3 [Fuerst et al., 1986]). The precise 5′ and 3′ ends of the mini-genomes are determined by the positioning of the T7 promoter and a hepatitis delta virus ribozyme (Wu et al., 1989), respectively. These mini-genomes are then assembled into nucleocapsids (with the SeV N protein), and transcribed into mRNA (and possibly replicated) via the SeV P-L polymerase proteins, provided from cotransfected pGEM plasmids (Curran et al., 1991). When the editing regions of the SeV or bPIV3 P genes were placed within the transcription unit of the mini-genome, the resulting mRNAs were found to contain G insertions whose distribution closely ressembled that found in their respective natural virus infections (Jacques et al., 1994). These results suggested that this system was behaving faithfully, and that the postulated counting mechanism was determined by *cis*-acting sequences contained within the 103-nt cassette containing the editing site. Unexpectedly, although the number of Gs in the run is conserved poorly among even closely related viruses (it varies from 3–7 Gs), changing the length of the G run in either the SeV or PIV3 editing regions led to both

null phenotypes and ones in which nucleotides were deleted rather than inserted (Jacques et al., 1994).

The template for paramyxovirus RNA synthesis is the viral nucleocapsid rather than naked RNA, where each nucleocapsid protein (N) subunit is associated with precisely 6 nt (Egelman et al., 1989). During the course of this work, the exact number of nucleotides in a SeV and measles virus genome chain was found to be important for the efficiency of genome amplification, especially when multiple rounds of replication are involved (Calain & Roux, 1993; Sidhu et al., 1995). The lengths of all natural SeV genomes (standard and defective interfering, DI, which can vary 20-fold in size), are, in fact, multiples of 6. Changing the length of natural DI genomes to ones that are not multiples of 6 was found to result in very inefficient genome amplification in transfected cells. This phenomenon, dubbed "the rule of six" (Calain & Roux, 1993), did not at first appear to exert the same effect on our synthetic mini-genomes used to study editing (Jacques et al., 1994). This paper reports that the precise genome length does, in fact, exert a major effect on RNA synthesis from our editing constructs, but in an unexpected way. When mini-genomes that are not of hexamer-length are used, antigenomes are generated in which nucleotides are inserted or deleted in the A_nG_n purine run, which readjust their length to multiples of six (genome length correction). Because hexamer-length antigenomes replicate efficiently, genomes are generated whose faithful transcription produces mRNAs with altered purine run lengths. The unusual requirement of paramyxoviruses for genomes of precise hexamer length may have evolved to maintain genome stability against insertions in the mRNA editing site during replication. The mRNA editing thus serves the dual and complementary function of maintaining hexamer genome length.

RESULTS

The most striking feature of the SeV transfected cell system is that it can reproduce the bPIV3 as well as the SeV mRNA editing pattern (or phenotype), depending on which editing region has been placed within the the mini-genome (cf Fig. 8, top). Further, when the regions upstream and downstream (defined relative to polymerase movement) of the SeV and bPIV3 editing regions were exchanged to create chimeric constructs, the editing pattern in both chimeras appeared to be determined mostly by the upstream regions (Jacques et al., 1994). Neither of the T7-generated genomes carrying the editing regions of the SeV and bPIV3 P genes were $6n + 0$ nt in length (i.e., a precise multiple of 6), and they would therefore not be expected to be amplified efficiently. The T7-generated genomes could, however, be copied directly into mRNA by the SeV polymerase, independent of their amplification by the SeV replication machinery. Transcription of mRNAs therefore did not appear to be sensitive to the precise length of the genome templates, at least in transfected cells.

Nevertheless, the finding that altering the SeV A_6G_3 purine run to A_6G_{1-5} (and that of bPIV3 from A_6G_4 to A_6G_{4-9}) could dramatically alter the insertion phenotypes (e.g., from insertions to deletions), was puzzling. We therefore examined the effect of changing the total number of nucleotides in the editing constructs directly, by simply filling in the upstream *Xba* I site or the downstream *Eco*R I site within the SeV A_6G_{1-5} series (located 25 nt upstream and 78 nt downstream of the editing site, respectively), adding 4 nt in each case per construct (Fig. 2). DNA-free mRNA (purified by pelleting through 5.7 M CsCl) was isolated from each transfection, copied into cDNA using oligo dT as a primer, and this was then amplified by PCR using as one primer an oligonucleotide complementary to the precise junction of the mRNA body and the polyA tail, so that predominantly SeV mini-mRNAs would be amplified. In this and subsequent experiments, we always withheld pGEM-L from some of the transfections as a negative control (L is the catalytic subunit of the viral polymerase). Only in those experiments in which these parallel samples failed to generate a detectable RT/PCR band were the remaining DNAs analyzed further; i.e., by limited primer extension to determine the number of purines in the purine run. In some cases, the number of Gs in the G run was also determined (using ddT) (see the Materials and methods, and Fig. 4). This method is not designed to quantitate the levels of mini-mRNAs produced between different transfections, but to determine the relative amounts of the various products present within a given transfection, and has been found to be reproducible in this latter respect.

The results are presented in Figure 2 as pairs of lanes for each construct. The left-hand lane of each pair shows the primer extension product obtained directly from the mini-genome plasmid used to initiate each transfection, and the position of the band here marks the unaltered length of the purine run. The right-hand lane shows the primer extension product obtained from the PCR-amplified transfected cell mRNA (5.7 M CsCl pellet RNA). When the results of Figure 2 are compared with those obtained with the unmodified A_6G_{1-5} constructs (cf. Fig. 3, Jacques et al., 1994), it is clear that the 4-nt insertions by themselves alter the phenotypes, and remarkably, this occurs in large part independently of where the insertion occurred. The A_6G_1 mini-genome, which had a null phenotype before insertion, changed to a robust −1 deletion phenotype upon insertion at either site ($G1/X^{+4}$ and $G1/R^{+4}$). The A_6G_5 mini-genome, where the −1 deletion phenotype was first noticed in the uninserted A_6G_{1-5} series, on the other hand, changed to the wild-type (50% fre-

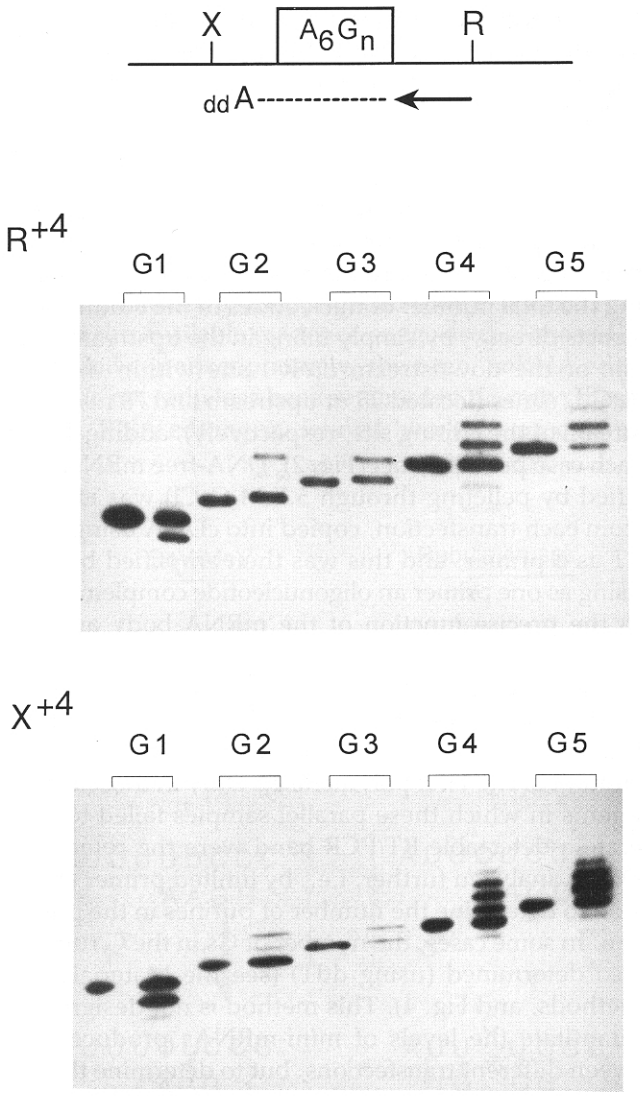

FIGURE 2. Effect of 4-nt insertions in the A_6G_{1-5} mini-genomes. Constructs expressing mini-genomes containing a run of 1–5 C residues (G1–G5), and in which either the upstream *Xba* I site or the downstream *Eco*R I site had been filled in (X^{+4} and R^{+4}, respectively), were transfected into cells that also synthesize the SeV N, P, and L proteins (see the Materials and methods). After 40 h incubation, cytoplasmic mRNA (5.7 M CsCl pellet RNA) was prepared and reverse transcribed into DNA with oligo dT as a primer. This DNA was then amplified by PCR using a SeV-polymerase-generated mRNA-specific primer (see the Materials and methods). Lengths of the purine runs in the amplified products were determined by extending a primer whose 3′ end abuts the purine run (horizontal arrow, above) in a reaction containing ddATP, to terminate the extension upstream of the purine run. The left-hand lane of each pair of lanes shows the primer extension product obtained directly from the mini-genome plasmid used for each transfection, and marks the unaltered length.

quency) +1 phenotype ($G5/R^{+4}$), or to a +1/+2 phenotype ($G5/X^{+4}$). Finally, the A_6G_4 mini-genome, which had a very attenuated +1 phenotype before insertion, changed to a robust +1/+2 (R^{+4}) or a +1/+2/+3 phenotype (X^{+4}) on adding four bases.

The requirement for hexamer genome length

The precise number of nucleotides of the genome template, and possibly also the number of Gs in the G run, appear to be important in determining the nature and extent of the insertions/deletions in the purine run of the mini-mRNAs. To determine the relative importance of the number of Gs in the G run per se, we constructed a series of mini-genomes with invariant A_6G_3, A_6G_4 (Fig. 3), or A_6G_5 (not shown) sequences, but in which the total number of nucleotides in the genome was varied by the insertion of 8–13 nt at a downstream *Eco*R V site (RV^{+n} series). The insertion/deletion phenotypes of

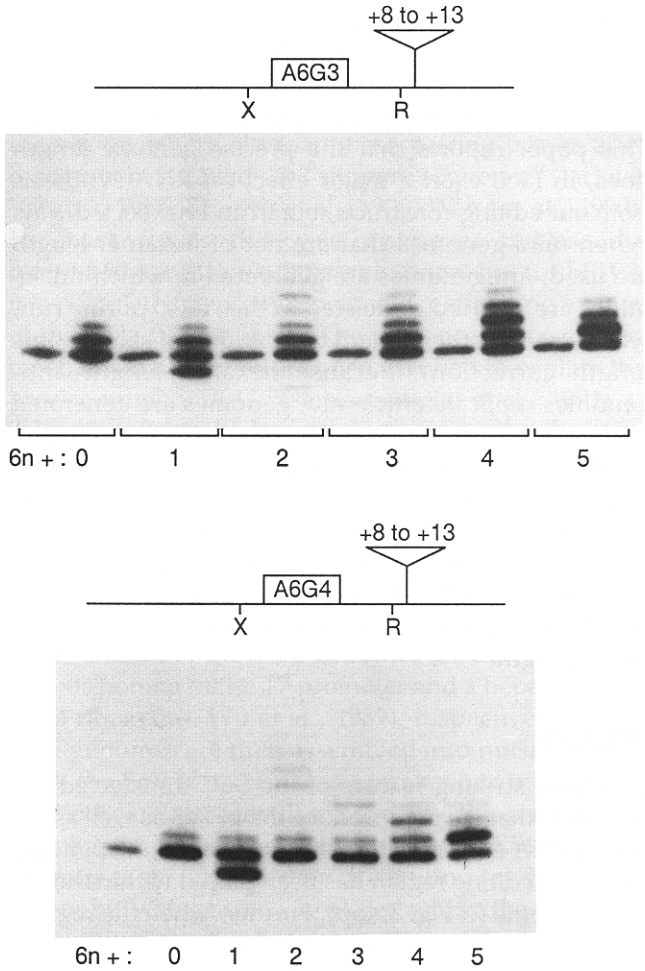

FIGURE 3. Effect of total chain length on the A_6G_3 and A_6G_4 mini-genomes. Mini-genomes containing either the A_6G_3 or A_6G_4 purine runs, and in which their total chain lengths were altered by the insertion of 8–13 nt at a downstream *Eco*R V site (elevated triangle, above) to generate mini-genomes that were $6n + 0$ to $6n + 5$ long (indicated below each panel), were transfected into cells. Lengths of the purine runs of the resulting mini-mRNAs were determined by primer extension in the presence of ddATP, as described for Figure 2. The left-hand lane of each pair of lanes (top panel) shows the primer extension product obtained directly from the mini-genome plasmid used for each transfection. The left-most lane in the bottom panel shows the primer extension product obtained directly from the $6n + 0$ mini-genome plasmid.

these series of mini-genomes are shown in Figure 3, where the samples are displayed according to their total length with respect to the hexamer rule, i.e., $6n+0$, $6n+1$, $6n+2$, etc. For reference, the A_6G_{1-5}/X^{+4} and $/R^{+4}$ series of Figure 2 are also in the order $6n+1$–5. The A_6G_{3-5}/RV^{+n} series contains 103 nt of the wild-type P gene (nt 1028–1130; the third G of the run is nt 1053, cf. Fig. 8, top), and this sequence plus the 1G and 2G insertions, respectively. The insertion/deletion phenotypes of their mini-mRNAs (Fig. 3) were found to be similar to each other and the A_6G_{1-5}/X^{+4} and $/R^{+4}$ series (Fig. 2) in several respects. For example, the most striking of the new phenotypes, the −1 deletion, occurs each time and only when the mini-genome is $6n+1$ nt long, even though the number of Gs in the G run within this group varies from 1 to 5. The $6n+1$ genome length then appears to be both necessary and sufficient to specify the single-base deletion. Two other phenotypes also appear to be determined predominantly by the length of the mini-genome: (1) the wild-type SeV phenotype (+1 at ca. 50% frequency, with +2 at a much reduced frequency) is always found when the mini-genome is $6n+5$ nt long, and (2) the phenotype in which the +2 band is at least as abundant as the +1 band is somehow linked to a mini-genome length of $6n+4$. A variation of this phenotype, (+1/+2/+3 at equal frequencies) occurs in the A_6G_4/X^{+4} construct (Fig. 2). The original A_6G_{1-5} series ($6n+3$, $6n+4$, $6n+5$, $6n$, and $6n+1$ long, respectively (Fig. 3; Jacques et al., 1994), and the A_6G_5/RV^{+n} series (not shown), are also in perfect accord with these observations. It is thus evident that there is an overall pattern of length changes in the purine run of the mini-mRNAs that is determined by the total number of nucleotides in the mini-genome according to a hexamer rule, and that is independent of the number of Gs in the G run.

Each of the A_6G_{3-5}/RV^{+n} series contains a construct whose length is a multiple of 6, and it is here that the differences between these series are most evident. The wild-type $A_6G_3^{6n+0}$ construct inserts a single base within this sequence at ca. 30% frequency (top, Fig. 3), and this frequency of insertion is maintained (or increased) throughout the series. That which contains the A_6G_4 sequence inserts a single base within this sequence at ca. 10% frequency (bottom, Fig. 3), and this frequency of insertion is also maintained (or increased) throughout the series. However, that which contains the A_6G_5 sequence does not insert a single base within this sequence at a detectable frequency (not shown). Thus, (1) the ability of constructs of hexamer genome length to generate mRNAs with insertions in the purine run is very sensitive to mutations in and around this sequence (see below), and (2) when such insertions occur in the $6n+0$ construct, they continue in constructs of the same series that are not of hexamer length, i.e., their presence is unaffected by the precise genome length of the remainder of the series.

The complexity of the insertion/deletion phenotypes

mRNA editing in natural infections occurs by expansion of the G run within the A_nG_n purine run. However, both A deletions and insertions, although very rare, were noted in P gene mRNAs made from natural SeV templates (Vidal et al., 1990b), and the phenotype that consists of a ladder of bands starting at position +4 and extending ca. 12 nt to the top of the gel (cf. A_6G_3/RV^{6n+2}, Fig. 4) is reminiscent of polyadenylation. The phenotypes described so far were determined by limited primer extension with ddATP, which is incorporated upstream of the purine run and determines the sum of the A and G changes. By also examining the ddTTP primer extension pattern (which determines only the G changes), one can estimate the fraction of insertions/deletions that are due to A or G changes in each sample.

The wild-type SeV +1 phenotype, associated with $6n+5$ genomes, was found to be due to a G insertion because the pattern of bands is basically the same when ddATP or ddTTP is used to limit the extension

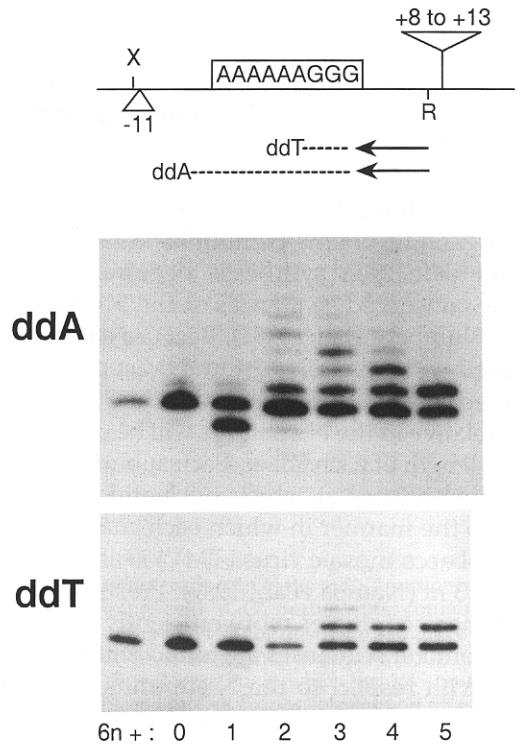

FIGURE 4. Effect of chain length of A_6G_3 mini-genomes with an upstream 11-nt deletion on both A and G changes in the purine run. This experiment is similar to that described in Figure 3, except for the mini-genome series, which contained an 11-nt upstream deletion, and the fact that primer extension in the presence of ddTTP, which terminates the extension opposite the last A of the A run (schematized above), was also conducted (bottom panel). The left-most lane in each panel shows the primer extension product obtained directly from the $6n+0$ mini-genome plasmid.

(Fig. 4). The −1 deletion phenotype of the $6n + 1$ mini-genomes, in contrast, was predominantly due to an A deletion because the number of Gs in the G run does not change under conditions where ca. half of the mini-mRNAs have deleted 1 nt (Fig. 4). That an A rather than a G residue is (sometimes) deleted in this phenotype was not unexpected, because −1 nt deletions occur as well with A_6G_1 slippery sequences (Fig. 2). A single G cannot be deleted in the stuttering model that depends on a realignment of the nascent chain downstream (Fig. 1), where the base pairing possibilities no longer exist. More complex editing phenotypes in which both A and G changes take place also occur. The $(6n + 4)$-associated phenotype, where +1 and +2 changes predominate and where the +2 frequency is equal to (or greater than) the +1 event, appears to be composed mostly of mini-mRNAs in which either a single G has been added, or ones in which both an A and a G were added (Fig. 4). This complexity extends to the "polyA" phenotypes observed with the $6n$ +2/3 mini-mRNAs.

The insertion/deletion phenotypes and hexamer genome length

The template for paramyxovirus RNA synthesis is the N:RNA nucleocapsid in which each N subunit is associated with precisely 6 nt, and the exact number of nucleotides in a SeV genome chain is important for the efficiency of genome amplification (Egelman et al., 1989; Calain & Roux, 1993). A working model for how the hexamer rule affects the efficiency of genome replication is outlined in Figure 5. Assembly of genome and antigenome chains is thought to take place concomitant with their synthesis, beginning and flush with the conserved sequences at the 5′ ends of these chains (Blumberg et al., 1981). Because each subunit is associated with precisely 6 nt, the exact position of the conserved (promoter) sequences at the 3′ ends of these chains relative to the N subunits will be determined by the total length of each chain. Each subunit is proposed to contain six sites at which nucleotides are bound, similar to the manner in which each coat protein subunit of tobacco mosaic virus (TMV) is associated with precisely 3 nt (Namba et al., 1989). Presumably, the viral polymerase initiates more efficiently when the *cis*-acting promoter sequences are found within the correct context with respect to the N subunits, as shown in Figure 5. For example, we have found recently that the insertion of a block of 6 nt at positions 47 or 67 (near the 3′ end of the genome) is well tolerated with respect to DI genome amplification efficiency, but that the addition of a total of 6 nt by the insertion of 2, 3, or 4 nt at position 47 and 4, 3, or 2 nt at position 67, respectively, is very poorly tolerated. The relative inactivity of these latter constructs is presumably due to the alteration of the hexamer phase of a promoter element

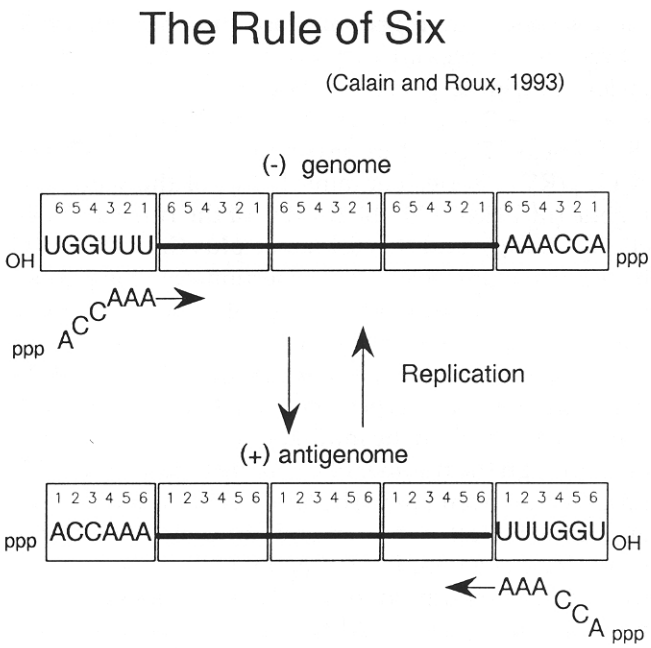

FIGURE 5. Paramyxovirus genome replication and the requirement for hexamer genome length. Genome and antigenome nucleocapsids are shown as a linear array of N subunits (open rectangles), each with six sites for binding nucleotides (numbered 1–6, 5′ to 3′). Only the sequence of the six 5′ (ppp) and 3′ (OH) conserved bases are shown. Nucleocapsid assembly occurs concomitantly with synthesis, and is proposed to initiate flush with the 5′ $_{ppp}$ACCAAA hexanucleotide. Chains that are not $6n + 0$ long alter the position of the 3′ $_{OH}$UGGUUU promoter element relative to the terminal N subunit and are therefore initiated inefficiently.

within positions 47–67 with respect to the N subunits (Pelet et al., 1996).

The notion of distinct nucleotide binding sites on N, and that of the phase of a *cis*-regulatory sequence relative to these sites, would apply to the mRNA editing site as well. The phase of the A_6G_3 sequence relative to the N subunits in the three A_6G_{3-5}/RV^{+n} series, for example, would be identical for each construct of these series because the *Eco*R V site used to vary the length lies between the editing site and the N assembly origin at the 5′ end of the genome template, and compensates for the G insertions. To determine whether the subunit phase of the purine run affected the insertion/deletion phenotype, we examined another series of constructs that contain the wild-type A_6G_3 sequence, but in which 11 nt of the upstream sequence was deleted. Because this 11-nt deletion does not lie in between the editing site and the N assembly origin at the 5′ end of the genome (like the *Eco*R V site), this results in the phase of the A_6G_3 sequence of each $6n + x$ construct in this series being shifted by 1 position toward the 3′ end of the genome (to the left as depicted in Fig. 5) relative to the undeleted A_6G_{3-5}/RV^{+n} series. When the mini-mRNAs of the $\Delta 11/A_6G_3/RV^{+n}$ series were examined (Fig. 4), however, they were found to be very similar in their insertion/deletion patterns to

those of the A_6G_{3-5}/RV^{+n} series (Fig. 3). The phase of the purine run relative to the N subunits therefore does not appear to be important in specifying the insertion/deletion phenotype. We also note that the $6n + 0$ construct of this series is again the only one of the six in which the length of the purine run remains unchanged (Fig. 4).

In a similar vein, we also examined a series of constructs that contained variable amounts of sequence immediately upstream of the A_6G_3 editing site, combined with various insertions in the downstream *Eco*R V site, such that all mini-genomes were $6n + 0$ long, but in which the A_6G_3 sequence was displaced into all six hypothetical N subunit phases (Fig. 6). The resulting mRNAs from all six mini-genomes, however, were found to be remarkably devoid of insertions and deletions (Fig. 6). The absence of insertions/deletions in the mini-mRNAs of these constructs (with upstream-deleted editing regions) is clearly associated with the precise hexamer length of the mini-genomes.

Genome replication and hexamer length

With the exception of the A_6G_3 and A_6G_4/RV^{+n} series, where the +1G insertion consistently occurs independent of genome length (Fig. 3), insertions and deletions in the mini-mRNAs of the remaining series (with more highly mutated editing regions) appeared to depend on the mini-genome length *not* being a multiple of six. The patterns of the $6n + 1$, $6n + 4$, and $6n + 5$ constructs, moreover, were curiously consistent in their movement to reestablish hexamer genome length, as is the presence of an enhanced band at +3 in the $6n + 3$ construct, and at +4 in the $6n + 2$ construct (Fig. 3). The question therefore arose as to whether these deletions and insertions also occurred during genome replication, and those that were now of hexamer length were being selectively amplified because of their increased ability to be replicated. Insertions and deletions would then appear in the mRNA population even for faithful mRNA synthesis.

Unlike unassembled mRNAs, which pellet in CsCl density gradients, genomes and antigenomes band at 1.31 g/mL, because their nucleocapsids are composed mostly (97% by weight) of N protein. In our system, genome nucleocapsids result from the assembly of chains made by either the T7 RNA polymerase (which does not detectably alter the length of the purine run, irrespective of the genome length [Fig. 6, and not shown]) or the SeV polymerase, and so examination of the CsCl-banded genomes would underestimate the frequency of changes introduced by the SeV enzyme. Antigenomes, however, are made only by the SeV polymerase. Unfortunately, one cannot use a SeV polymerase product-specific primer (i.e., one that spans the precise 3′ end of the mRNA and the polyA tail) in amplifying the mini-antigenomes. These amplifications are thus inherently more sensitive to DNA contamination, which is very extensive and varied due to the multiple transfected plasmids and the concomitant vaccinia virus infection (Garcin et al., 1995). For example, in contrast to the examination of the viral mRNAs, when L coexpression was withheld from some of the transfections to ensure that the changes in the purine run length were due to the SeV enzyme, a band at the zero position was nevertheless often present following PCR amplification. We found, however, that it was possible to examine directly the SeV-replicated antigenomes by limited primer extension, without PCR amplification, by (1) improving the rep-

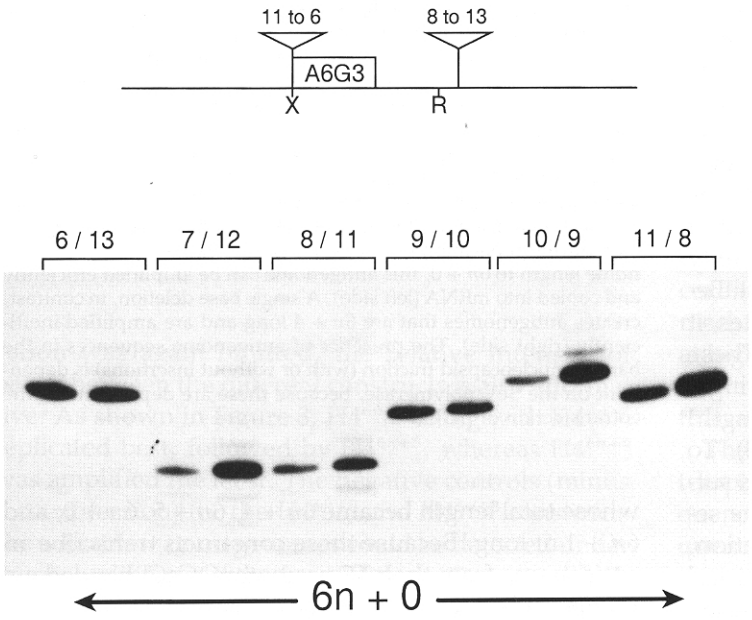

FIGURE 6. Effect of changing the A_6G_3-N subunit phase on the length of the mRNA purine run. Six mini-genomes with A_6G_3 purine runs, containing various amounts (6–11 nt) of upstream sequence and 8–13 nt inserted at the downstream *Eco*R V site, such that all are $6n + 0$ long, but in which the A_6G_3 purine run was displaced into all six possible subunit phases, were examined. The left-hand lane of each pair of lanes shows the primer extension product obtained directly from the mini-genome plasmid used for each transfection, and marks the unaltered length.

lication efficiency of our DI constructs (by the inclusion of viral sequences near the 5′ end of the genome, to generate the SH22 backbones [see the Materials and methods; Fig. 8, top]), (2) replacing the pGEM expressing the SeV N protein with one expressing the equally functional N protein from human parainfluenza virus type 1, to limit recombination between the pGEM-N and the editing constructs, and (3) deleting 30 nt just downstream of the A_6G_3 in the support plasmid expressing the P protein, which only slightly affects its function (Curran, 1996), but eliminates the binding site for the limited primer extension oligonucleotide from the pGEM-P transcripts. As shown in Figure 7, when the SH22^{6n+0} construct was amplified in the transfected cell system and the CsCl-banded antigenomes were examined directly, a single, strong band at the zero position was found, whose presence was entirely dependent on the coexpression of the viral L protein (lanes $6n + 0$, ±L). The $6n + 4$ and $6n + 5$ constructs, in contrast, each contained a second band, which was at least as intense as the zero band, at the +2 and +1 positions, respectively. These results support the view that a mechanism is operating to correct the hexamer length during antigenome synthesis. They also suggest that the deletions and insertions found in the mRNA population (in Figs. 2, 4) might not have occurred during mRNA synthesis, but during replication. Only the experiments presented in Figure 3, in which a single nucleotide is inserted into the purine run even when the editing construct is precisely of hexamer length, are likely to represent examples of insertion during mRNA synthesis, or bona fide mRNA editing as defined previously.

The experiments presented in Figures 4 and 6 suggest that constructs with deletions upstream of the purine run do not insert nucleotides during mRNA synthesis, because the +1 band is not generated here for constructs of hexamer length. They can, however, correct their genome length (Fig. 4), and this suggests that the *cis*-acting sequences that determine insertions during mRNA synthesis are more extensive than those that determine insertions/deletions during antigenome synthesis (genome length correction). The importance of hexamer length for efficient genome replication was discovered by altering the length of DI-H4, a natural 1,410-nt long DI genome that contains sequences only from the 5′ end of the nondefective genome (Fig. 8). Because DI-H4 does not contain mRNA initiation sites, it acts only as a template for genome replication. These H4 genomes were found previously to correct their genome length (when they are not of hexamer length) very inefficiently, if at all (Calain & Roux, 1993). To investigate whether the SeV purine run alone was sufficient to allow for genome length correction, the sequence A_6G_3 (in the normal plus-sense orientation, within a 28-nt long polylinker; see the Materials and methods) was placed in a series of H4 derivatives (Fig. 8),

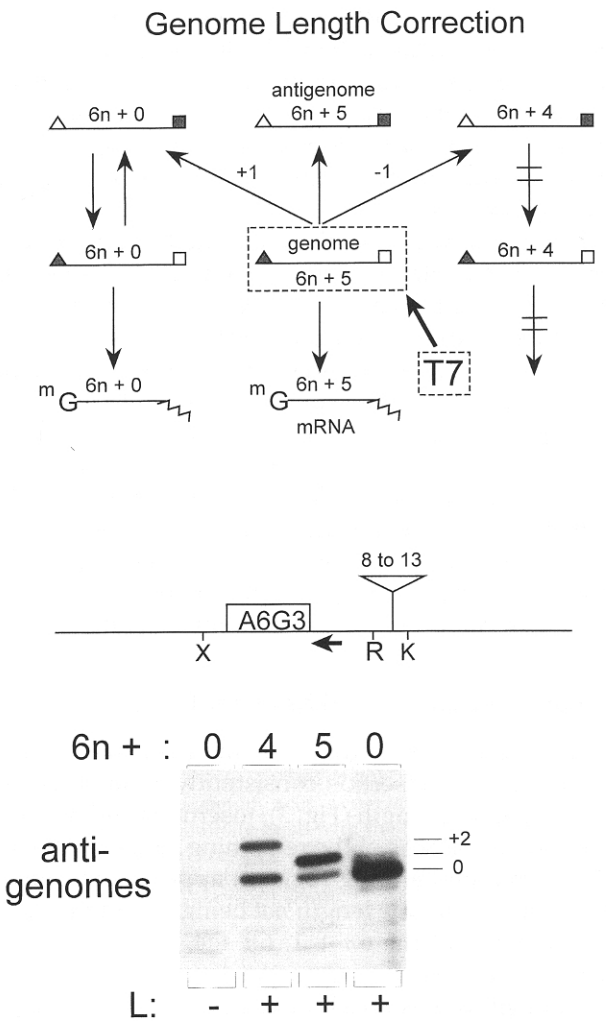

FIGURE 7. Genome length correction during antigenome synthesis. Several of the constructs carrying the wild-type A_6G_3 cassette within the SH22 backbone (cf. top of Fig. 8) were transfected into cells, as for Figure 3, but (1) pGEM-N^{SeV} was replaced with pGEM-N^{PIV1}, (2) pGEM-P was replaced with pGEM-P$^{\Delta 310\text{-}320}$ (see the Materials and methods), and (3) the antigenome RNA of the CsCl-banded nucleocapsids was examined directly by limited primer extension (see the Materials and methods). The distribution of the purine run lengths was determined as before in the presence of ddATP (bottom panel). As a negative control, pGEM-L was withheld from the $6n + 0$ transfection (indicated below). An interpretation of the events leading to the single base addition during replication of the $6n + 5$ genome is outlined above. Genomic and antigenomic promoters are shown as triangles and squares at their ends, respectively. Synthesis of the $6n + 5$ genome is largely due to the T7 RNA polymerase, and is shown in a dotted rectangle. When a single base addition occurs during replication by the SeV polymerase, which restores genome length to $6n + 0$, this antigenome can be amplified efficiently and copied into mRNA (left side). A single base deletion, in contrast, creates antigenomes that are $6n + 4$ long and are amplified inefficiently (right side). The presence of antigenome sequences in the banded nucleocapsid fraction (with or without insertions) is dependent on the SeV polymerase, because these are dependent on the cotransfection of pGEM-L.

whose total length became $6n + 4$, $6n + 5$, $6n + 0$, and $6n + 1$ nt long. Because these constructs transcribe an antigenome from their T7 promoters, CsCl-banded genomes were examined directly. Because PCR amplifi-

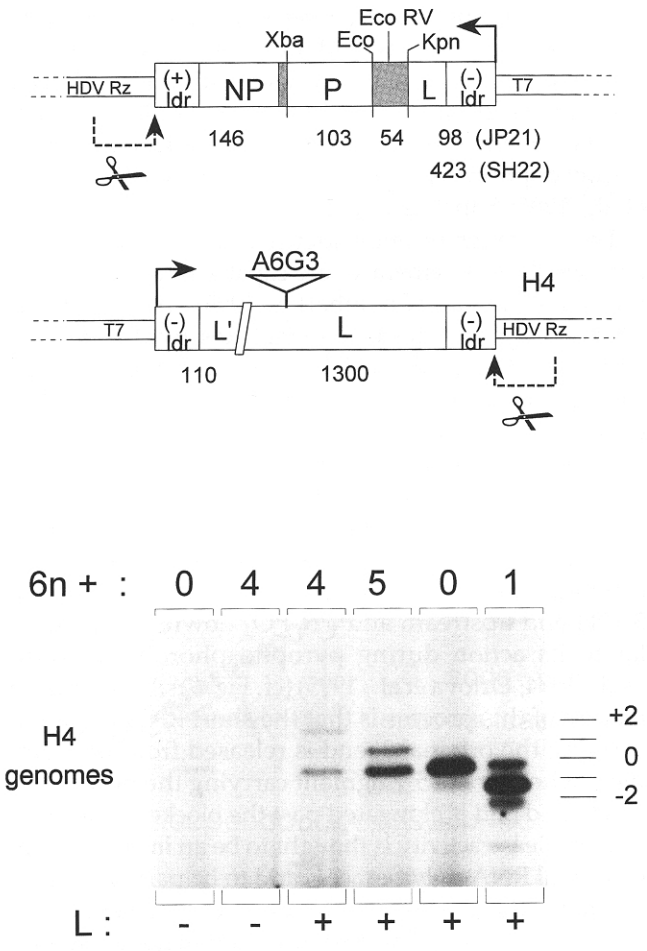

FIGURE 8. Genome length correction during replication of copy-back DI genomes. Organization of the various vectors used to study editing are outlined above. JP21 and SH22 are modeled on internal deletion DI genomes, and contain various amounts of both ends of the nondefective genome. The regions in grey represent polylinker sequences, and key restriction sites are indicated. Letters inside the boxes refer to viral genes, and numbers below indicate the length of each region. The copy-back DI-H4 is shown below, containing 1,300 nt of the 5′ end of the nondefective genome and an inverted repeat of the last 110 nt at the 3′ end. The position of the A_6G_3 insertion is indicated. Positions of the T7 promoters and HDV ribozymes are also indicated. Several of the H4 constructs carrying the A_6G_3 insertion were transfected into cells, as for Figure 7, but the genome RNA of the CsCl-banded nucleocapsids was examined directly by limited primer extension (see the Materials and Methods), and the distribution of the pyrimidine run lengths (of the [−] genome) was determined as before in the presence of ddATP (bottom panel). As a negative control, pGEM-L was withheld from two of the transfections (indicated below).

cation was again omitted, the relative intensity of bands between the different constructs is also informative. As shown in Figure 8, $H4^{6n+0}$ along with $H4^{6n+1}$ replicated best, followed by $H4^{6n+5}$, whereas $H4^{6n+4}$ was amplified the least. The negative controls (minus L coexpression) for the $H4^{6n+0}$ and $H4^{6n+4}$ constructs show that few or no genomes could be detected under these conditions. Of the four H4 derivatives amplified, only the length of the A_6G_3 run of $H4^{6n+0}$ remained unaltered. Those resulting from the $H4^{6n+4}$, $H4^{6n+5}$, and $H4^{6n+1}$ amplifications now also contained prominent bands representing genomes with 2 and 1 additional nucleotide, and 1 less nucleotide, respectively, i.e., whose hexamer length had been corrected by the shortest route. The $H4^{6n+1}$ amplification was exceptional, in that the corrected genome was now the most abundant species (62%). We also note a prominent band here at position −2, which is almost as intense as that at the zero position, and which is not of hexamer length because it has deleted one too many nucleotides. Taken together, these results show that copy-back DI genomes containing the P gene slippery sequence alone can correct their genome length almost as efficiently as internal deletion DI genomes containing 103 nt of the P gene sequence.

DISCUSSION

Paramyxovirus polymerase stuttering (pseudo-templated transcription) is thought to form the polyA tails of the viral mRNAs, and to alter the P gene mRNAs by the controlled expansion of a short G run. The term mRNA editing is used to describe this latter phenomenon, because of the programmed translational consequences of the insertions, and because the length of the genome G run in natural infections remains unaltered. In the most extensive study (Vidal et al., 1990a), >300 clones containing the insertion region of the SeV genome were examined, without finding a single case of insertions or deletions. This work, however, provides evidence that pseudo-templated transcription can also occur at the same purine run used for G insertions during mRNA synthesis, when the viral polymerase is copying the genome template during antigenome synthesis. This latter phenomenon (genome length correction) is different from the G insertions during mRNA synthesis in three important respects: (1) it can delete as well as insert purines (and A as well as G residues); (2) it can be found at a significant frequency only when the genome is not of hexamer length; and (3) only the A_6G_3 sequence itself is required, whereas the G insertions during mRNA synthesis require additional upstream sequences. Given that genomes that are precise multiples of six are replicated preferentially, this leads to a mechanism that corrects the genome length according to this rule. Our results suggest a two-step mechanism in which imperfect length correction occurs at high frequency during antigenome synthesis, and is then refined by the preferential replication of templates that are of precise hexamer length.

The frequency of G-inserted mRNAs arising from templates of hexamer length directly reflects the frequency of the pseudo-templated transcription. The frequency of A or G-inserted antigenomes arising from templates not of hexamer length, however, is also conditioned by their relative efficiencies in replication and

the number of rounds of replication that have occurred. The relative absence of +1 bands in the amplification of the SH22/$A_6G_3^{6n+4}$ and $H4^{6n+4}$ templates, and that of +2 bands in the corresponding $6n + 5$ constructs (Figs. 7, 8), suggests that multiple rounds of amplification have occurred, hence the frequency of the pseudo-templated synthesis may be obscured. We have pointed out the unexpected −2 band resulting from amplification of the $H4^{6n+1}$ genome because it would not be amplified selectively. This −2 band is even more prominent in transfections of internal deletion $6n + 1$ DI constructs carrying the bPIV3 editing region (not shown), but the −2 band is absent in the majority of our $6n + 1$ constructs of Figures 2, 3, and 4. The reasons for the variable appearance of these nonhexamer-length bands in the antigenome population is unclear. However, because this −2 band is not expected to be amplified selectively, these results suggest that under some circumstances pseudo-templated RNA synthesis occurs during genome replication at significant rates. Moreover, it is not impossible that pseudo-templated replication occurs at rates approaching those in mRNA synthesis even for genomes of hexamer length, but these changes are not seen because nonhexamer-length antigenomes are not copied back into genomes, and/or are less stable than those of hexamer length. Further insight into this mechanism may be obtained if genome replication can be restricted to a single round.

SeV genome length correction appears to occur predominantly (or exclusively) in the A_6G_3 purine run used for mRNA editing, as opposed to that which is expanded in polyadenylation ($ANTAAGA_5$). For example, copy-back DI-H4 genomes that contain polyA sites but not editing sites are amplified only very poorly when not of hexamer length (Calain & Roux, 1993), but largely escape this requirement when an editing site slippery sequence is introduced (Fig. 8). Further, all negative-strand virus RNA polymerases that polyadenylate their mRNAs are thought to do so by stuttering on a short run of template U residues (4–7 nt long). They are thus all capable of pseudo-templated transcription, but only paramyxovirus (and Ebola filovirus, see below) polymerases are known to edit their mRNAs, and only paramyxoviruses are known to require genomes of a precise integer length for efficient replication. The phenomena of mRNA editing and genome length correction thus appear to be linked, and distinct from that responsible for polyadenylation. In other mononegaloviruses, such as the rhabdovirus VSV or the pneumovirus RSV, for which mRNA editing has not been described, genome replication does not appear to be governed by an integer rule (Pattnaik et al., 1995; Samal & Collins, 1996). The hexamer rule may thus be required to maintain the length of the P gene purine run against the instability that occurs during genome replication. The proposed requirement for a given phase of the promoter relative to N subunits for efficient initiation would then have coevolved with the ability of these polymerases to edit their mRNAs. If so, the replication efficiency of Ebola filovirus may also be governed by an integer rule, because this virus cotranscriptionally edits its G gene mRNA (Volchkov et al., 1995; Sanchez et al., 1996).

There is much recent evidence that bacterial and eukaryotic RNA polymerases are capable of extraordinary internal flexibility (Chamberlin, 1995). When cellular RNA polymerase ternary complexes (with template and nascent RNA) are stalled or arrested, their catalytic sites can be displaced upstream along the nascent RNA chain by as much as 17 nt without disrupting the ternary complex (Johnson & Chamberlin, 1994; Nudler et al., 1994). In this altered state, the catalytic site can be induced to act as an endonuclease (by gre A/B in *Escherichia coli*, and TFIIS in yeast) by adding water across an internal phosphodiester bond to generate a 3′ OH end upstream and a 5′ PO_4 downstream, similar to its action during pyrophosphorolysis (Rudd et al., 1994; Orlova et al., 1995) (cf. Fig. 9). A remarkable feature of this process is that the short RNA fragment carrying the original 3′ end is released from the complex, whereas the 5′ fragment carrying the new 3′ end is retained and is elongated past the blockage site. This endonuclease activity is thought to be an intrinsic property of all RNA polymerases, and to be part of a mechanism for the processivity required of these enzymes. Two models for paramyxovirus polymerase pseudotemplated transcription are now plausible. In both, the polymerase pauses, and the nascent chain somehow disengages from the template and then realigns using alternate base pairing possibilities, such that a nucleotide will be added (or lost) when synthesis restarts (Fig. 1). In our original model, the nascent chain disengages without being cut, simply because there are only a few base pairs holding it in place. However, we know of no precedent for this during elongation. It is also possible that the SeV polymerase behaves like a cellular RNA polymerase encountering an elongation block (Fig. 9), and its catalytic site is displaced upstream and induced to cut the nascent chain, within the purine run. The new 3′ end must then realign on the template to be elongated, and the alternate base pairing possibilities of this slippery sequence leads to the insertions/deletions (Figs. 1, 9). Viral and cellular RNA polymerases have a propensity to slip back on the template during initiation while retaining the nascent chain, causing repetitions at the 5′ end of the RNA transcript (reviewed in Garcin et al., 1995). *E. coli* RNA polymerase does this spontaneously on certain promoters, a phenomenon called "primer shifting" (Borukhov et al., 1993; Severinov & Goldfarb, 1994). Primer shifting is also simulated by gre A/B, hence the endonuclease and slippage activities may be linked for the bacterial enzyme (Feng et al., 1994). There is thus ample precedence for

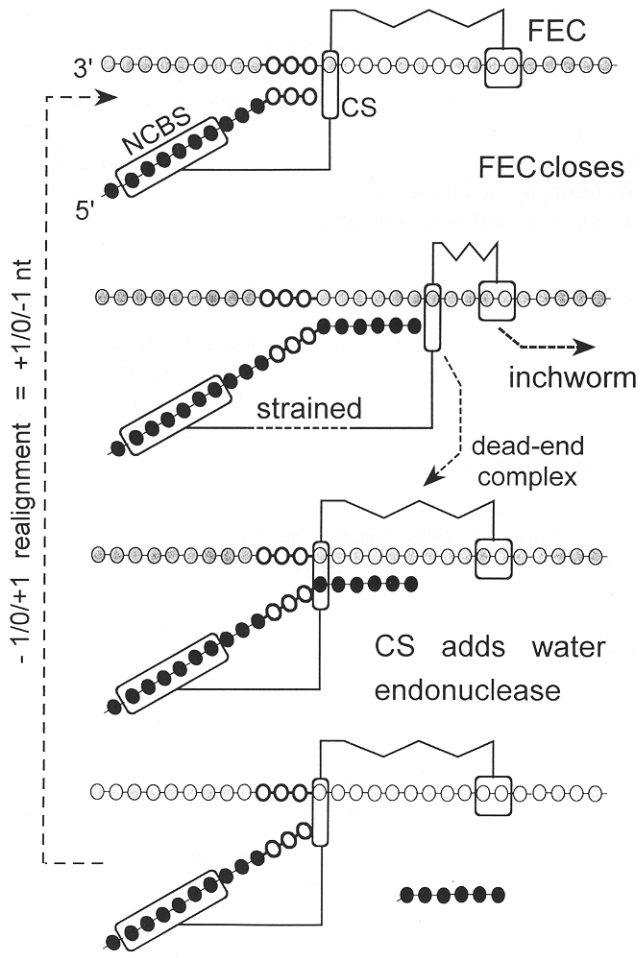

FIGURE 9. A cleavage/re-elongation model for SeV polymerase pseudo-templated RNA synthesis. The model for paramyxovirus cotranscriptional insertions (and deletions) is adapted from those of Chamberlin (1995) and Nudler et al. (1995) for cellular RNA polymerase elongation. Viral polymerase is shown as three interconnected rectangles, representing (from left to right) a nascent chain-binding site (NCBS, also called the tight product binding site), the catalytic site (CS), and the "front-end clamp" (FEC, a template-binding site). The template is shown above (3′ to 5′) and the nascent chain shown below (5′ to 3′), with only three base pairs formed between them. The ternary complex is first shown at the start of an "inch-worming" cycle, in which the FEC closes. Nucleotide addition continues when the FEC is closed, so that the CS approaches the FEC (second level). In normal inchworming, the NCBS is first open and follows the CS. It then closes, and the FEC opens and leaps ahead, completing the inchworm cycle. For paramyxoviruses, the model proposes that when the polymerase encounters the *cis*-acting editing signal, both the FEC and NCBS close (second level). Continued elongation causes the distance between the NCBS and the CS to lengthen, creating strain, which is eventually relieved by the disengagement of the CS from the 3′ end of the nascent chain and its repositioning upstream along the nascent chain, creating the dead-end complex (third level). The resulting dead-end complex is resolved by the CS adding water across an internal phosphodiester bond (which for paramyxoviruses is postulated to be within the A_nG_n purine run/slippery sequence), and the release of the original 3′ end fragment (bottom level). Precise realignment of the new 3′ end on the template (the only possibility for heteropolymeric sequences) leads to the faithful continuation of elongation. For "slippery sequences" such as A_nG_n (Fig. 1), alternate realignments can occur, determined by the base pairing possibilities (and possibly also residual strain in the polymerase), leading to insertions and deletions on re-elongation.

this latter cleavage model for paramyxovirus pseudo-templated RNA synthesis.

MATERIALS AND METHODS

Construction of mini-genomes

Construction of the internal deletion SeV mini-genome (JP21) is described in Jacques et al. (1994). Briefly, pJP21 is based on pSP65 with the mini-genomes inserted directly downstream of the T7 promoter, and carrying the hepatitis delta virus genomic ribozyme directly after the mini-genome (Fig. 8). Each JP21 mini-genome contains 98 nt of the 5′ (−) leader/L gene region of SeV, a 54-nt polylinker region into which the editing cassette 103 nt of SeV has been inserted between *Eco*R I and *Xba* I, and 146 nt of the 3′ end of the SeV (−) genome including the (+) leader/N gene region. SH22 derivatives are similar to JP21, but contain 325 nt more in the L gene (Fig. 8).

The $A_6G_{1\text{-}5}/R^{+4}$ and $/X^{+4}$ series were obtained by digesting and filling in the corresponding *Eco*R I and *Xba* I sites. The $A_6G_{1\text{-}5}/RV^{6n+x}$ mutants (+8 to +13) were constructed by inserting short palindromic oligonucleotides into the *Eco*R V site. The self-complementary oligonucleotides used were: add8 (5′-AGGGCCCT-3′), add9 (5′-ACGGACCGT-3′), add10 (5′-TAGGGCCCTA-3′), add11 (5′-TACGGTCCGTA-3′), add12 (5′-CATCCCGGGATG-3′), and add13 (5′-CTACGGTCCGTAG-3′).

The −11, −10, −9, and −6 deletions upstream from the SeV editing site were constructed by PCR mutagenesis using the following primers: short0 (5′-GCTCTAGACAAAAAAGGG-3′), short1 (5′-GCTCTAGATCAAAAAAGGG-3′), short2 (5′-GCTCTAGACTCAAAAAAGGG-3′), and short5 (5′-GCTCTAGACAAAAAGGG-3′). The −8 and −7 deletions were derived from the −6 and −11 deletions by filling in the *Xba* I site. The combination of *Eco*R V mutants (+8 to +13) with the upstream −11 deletion (Fig. 4) was constructed by subcloning the *Eco*R I-*Ssp* I 1.3-kb fragment of the Δ11 mutant into the various complementary *Eco*R I-*Ssp* I 2.2-kb fragments of the +8 to +13 series. The SeV series of hexamer-length compensatory mutations 6/13, 7/12, 8/11, 9/10, 10/9, and 11/8 of Figure 6 were constructed by subcloning the various *Eco*R I-*Ssp* I 1.3-kb fragments with the −11 to −6 upstream deletions into the complementary *Eco*R I-*Ssp* I 2.2-kb fragment with the +8 to +13 downstream insertions.

Construction of the copy-back DIH4 genomes in different phases are described in Calain and Roux (1993). The slippery sequence (A_6G_3 alone) in $6n$, $6n + 1$, $6n + 4$, $6n + 5$ phases were derived from psvDIH4/Rbz and its derivatives by inserting 28 nt (5′-GCTCTAG**AAAAAAGGG**TACCGAATTCCG) into the *Stu* I site (Fig. 8).

Sequences of all constructs were confirmed after subcloning.

Mini-genome expression and selective amplification of the mini-mRNAs

The various mini-genomes and the SeV N, P, and L genes were expressed in HeLa cells basically as described previously (Curran et al., 1991; Jacques et al., 1994). The cells were grown as monolayers in 9-cm dishes and infected with 2–5 pfu/cell of a vaccinia virus recombinant expressing T7 RNA polymerase (vTF7-3 [Fuerst et al., 1986]). At 1-h post infection, the medium was replaced with a transfection mix com-

posed of 20 μL of homemade transfectACE (Rose et al., 1991), pGEM-L (1 μg), pGEM-N (2.5 μg), pGEM-P (2.5 μg), and MEM up to 1 mL. After 2 h at 33 °C, an extra 6 mL of MEM were added and the cells were incubated for 40 h before harvesting. After removal of medium, the cells were solubilized and scraped into 150 mM NaCl, 50 mM Tris, pH 7.4, 10 mM EDTA, and 0.6% NP 40. Nuclei were removed by pelleting at 12,000 × *g* for 5 min. mRNA was recovered by pelleting the material (38,000 rpm, overnight, in an SW55 rotor) in a step gradient composed of a 5.7 M CsCl cushion (to remove DNA) and a 40% CsCl solution (to remove viral nucleocapsids). The RNA pellet was resuspended in TE and recovered by ethanol precipitation. A 1/4 sample of that solution was then treated with 10 units of RNase-free DNAse I (Boehringer). After 30 min at 37 °C, the DNase I was inhibited with 20 mM EDTA, the material was phenol-chloroform extracted, and then ethanol precipitated in the presence of 10 ng oligo $d(T)_{15}$.

After resuspension in reverse-transcriptase buffer according to the supplier's instructions (500 mM dNTP, 40 units RNAsin [Promega]), the polyadenylated mRNA was copied into single-stranded cDNA with 200 units MMLV-reverse transcriptase (Gibco-BRL) for 1–2 h at 42 °C. The product of reverse transcription was then extracted with phenol-chloroform and ethanol precipitated. A 1/20 aliquot was used for a hot-start touch-down PCR with 240 ng of each primer, (+) 191 primer (5′-GACCCTTTGCTTTGCTGCC), and $L(T)_{22}$ primer [5′-$(T)_{22}$CTTACT], in 10 mM Tris, pH 8.3, 50 mM KCl, 2 mM $MgCl_2$, 100 μM dNTP. The PCR mix was assembled at 80–90 °C to avoid nonspecific annealing of primers, and PCR was conducted in 50 μL with 1 unit Taq polymerase in a GeneAmp PCR system 9600 as follows; a preliminary denaturation was done for 2 min at 94 °C; then 8 cycles of touch-down PCR, 94 °C for 20 s, decreasing the annealing temperature by 1° from 67 to 60 °C at each cycle for 20 s, elongation at 72 °C for 20 s; then an extra 20–25 cycles, 94 °C for 20 s, 60 °C for 20 s, 72 °C for 20 s.

Limited primer extension and sequencing reactions

After gel purification on a 2% agarose gel, the PCR products were annealed to ^{32}P-primers (SeV-edit) complementary to the sequence immediately downstream from the editing site. Primer extension was performed in 10 μL at 37 °C for 6 min with 1 unit of T7 DNA polymerase (Pharmacia), in presence of 40 μM dGTP, dTTP, dCTP, and 4 μM dideoxy-ATP. Then 300 μM dNTP was added and the mix was incubated an extra 2 min to chase stalled complexes. The reaction was stopped by adding 4 μL STOP solution (95% formamide, 20 mM EDTA, 0.1% bromophenol blue, and xylene cyanol FF). Products were boiled 1 min and loaded on a 12.5% denaturing polyacrylamide gel. For the DNA markers, the same procedure was used starting from 1 μg plasmid DNA. For Figure 4, ddATP was substituted with ddTTP.

Examination of nucleocapsid RNA

The cytoplasmic extracts, prepared as described above, were centrifuged on 20–40% (w/w) CsCl density gradients to purify the viral nucleocapsids (Curran et al., 1991). RNA from purified nucleocapsids from the SH22 and DIH4 transfections were used directly for limited primer extension with, respectively, primers SeV-edit (5′-GATGTGTTCTCTCCTATG) and DIH4 edit H4+ (5′-GCTCGATAGGGCTCTAG), as described in Pelet et al. (1991). To further enhance the specificity of this direct examination, pGEM-N from human parainfluenza virus type 1 was transfected instead of that of SeV to limit recombination between pGEM-N and the editing constructs. Moreover pGEM-$P^{\Delta 310-320}$ was transfected instead of pGEM-P to eliminate the binding of oligonucleotides to pGEM-P transcripts that had been encapsidated nonspecifically.

ACKNOWLEDGMENTS

This work was supported by grants from the Swiss National Science Fund and the Human Frontiers Science Program Organization. We thank Laurent Roux, Joe Curran, and Dominique Garcin (Geneva), and Michael Chamberlin (Berkeley) for fruitful discussions.

Received June 21, 1996; returned for revision July 9, 1996; revised manuscript received August 1, 1996

REFERENCES

Benne R. 1993. Paramyxovirus P gene mRNA Editing. In: Benne R, ed. *RNA editing*. England: Ellis Horwood. pp 105–123.

Blumberg BM, Leppert M, Kolakofsky D. 1981. Interaction of VSV leader RNA and nucleocapsid protein may control VSV genome replication. *Cell 23*:837–845.

Borukhov S, Sagitov V, Josaitis CA, Gourse RL, Goldfarb A. 1993. Two modes of transcription initiation in vitro at the rrnB P1 promoter of *E. coli*. *J Biol Chem 268*:23477–23482.

Brierley I, Digard P, Inglis SC. 1989. Characterization of an efficient Coronavirus ribosomal frameshifting signal: Requirement for an RNA pseudoknot. *Cell 57*:537–547.

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.

Cattaneo R. 1991. Different types of mRNA editing. *Annu Rev Genetics 25*:71–88.

Chamberlin MJ. 1995. New models of transcription elongation and termination. *Harvey Lect 88*:1–21.

Curran J. 1996. Reexamination of the Sendai virus P protein domains required for RNA synthesis; a possible supplemental role for P. *Virology*. Forthcoming.

Curran J, Boeck R, Kolakofsky D. 1991. The Sendai virus P gene expresses both an essential protein and an inhibitor of RNA synthesis by shuffling modules via mRNA editing. *EMBO J 10*: 3079–3085.

Curran J, Kolakofsky D. 1990. Sendai virus P gene produces multiple proteins from overlapping open reading frames. *Enzyme 44*: 244–249.

Egelman EH, Wu SS, Amrein M, Portner A, Murti G. 1989. The Sendai virus nucleocapsid exists in at least four different helical states. *J Virol 63*:2233–2243.

Feng G, Lee DN, Wang D, Chan CL, Landick R. 1994. GreA-induced transcript cleavage in transcription complexes containing *E. coli* RNA polymerase is controlled by multiple factors, including nascent transcript location and structure. *J Biol Chem 269*:22282–22294.

Fuerst TR, Niles EG, Studier FW, Moss B. 1986. Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proc Natl Acad Sci USA 83*:8122–8126.

Garcin D, Iezzi M, Dobbs M, Elliott RM, Schmaljohn C, Yong Kang C, Kolakofsky D. 1995. The 5′ ends of Hantaan virus (Bunyaviridae) RNAs suggest a prime-and-realign mechanism for the initiation of RNA synthesis. *J Virol 69*:5754–5762.

Jacks T, Varmus HE. 1985. Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. *Science 230*:1237–1242.

Jacques JP, Hausmann S, Kolakofsky D. 1994. Paramyxovirus mRNA editing leads to G deletions as well as insertions. *EMBO J 13*: 5496–5503.

Jacques JP, Kolakofsky D. 1991. Pseudo-templated transcription in prokaryotic and eukaryotic organisms. *Gene & Dev 5*:707–713.

Johnson TL, Chamberlin MJ. 1994. Complexes of yeast RNA polymerase II and RNA are substrates for TFIIS-induced RNA cleavage. *Cell 77*:217–224.

Kolakofsky D, Curran J, Pelet T, Jacques JP. 1993. Paramyxovirus P gene mRNA editing. In: Benne R, ed. *RNA editing*. England: Ellis Horwood.

Lamb RA, Kolakofsky D. 1996. Paramyxoviridae: The viruses and their replication. In: Fields B, Knipe D, Howley P, eds. *Fields virology, 3rd ed*. New York: Raven Press. pp 1177–1204.

Namba K, Pattanayek R, Stubbs G. 1989. Visualization of protein-nucleic acid interactions in a virus. *J Mol Biol 208*:307–325.

Nudler E, Goldfarb A, Kashlev M. 1994. Discontinuous mechanism of transcription elongation. *Science 265*:793–786.

Nudler E, Kashlev M, Nikiforov V, Goldfarb A. 1995. Coupling between transcription termination and RNA polymerase inchworming. *Cell 81*:451–357.

Orlova M, Newlands J, Das A, Goldfarb A, Borukhov S. 1995. Intrinsic transcript cleavage activity of RNA polymerase. *Proc Natl Acad Sci USA 92*:4596–4600.

Pattnaik AK, Ball LA, LeGrone A, Wertz GW. 1995. The termini of VSV DI particle RNAs are sufficient to signal RNA encapsidation, replication, and budding to generate infectious particles. *Virology 206*:760–764.

Pelet T, Curran J, Kolakofsky D. 1991. The P gene of bovine parainfluenza virus 3 expresses all three reading frames from a single mRNA editing site. *EMBO J 10*:443–448.

Pelet T, Delenda C, Gubbay O, Garcin D, Kolakofsky D. 1996. Partial characterization of a Sendai virus replication promoter and the rule of six. *Virology*. Forthcoming.

Rudd MD, Izban MG, Luse DS. 1994. The active site of RNA polymerase II participates in transcript cleavage within arrested ternary complexes. *Proc Natl Acad Sci USA 91*:8057–8061.

Samal SK, Collins PL. 1996. RNA replication by a respiratory syncitial 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.

Sanchez A, Trappier SG, Mahy BWJ, Peters CJ, Nichol ST. 1996. The spike glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc Natl Acad Sci USA*. Forthcoming.

Severinov K, Goldfarb A. 1994. Topology of the product binding site in RNA polymerase revealed by transcript slippage at the phage λ P_L promoter. *J Biol Chem 269*:31701–31705.

Sidhu MS, Chan J, Kaelin K, Spielhofer P, Radecke F, Schneider H, Masurekar M, Dowling PC, Billeter MA, Udem SA. 1995. Rescue of synthetic measles virus minireplicons; measles genomic termini direct efficient expression and propagation of a reporter gene. *Virology 208*:800–807.

Vidal S, Curran J, Kolakofsky D. 1990a. Editing of the Sendai virus, P/C mRNA by G insertion occurs during mRNA synthesis via a virus-encoded activity. *J Virol 64*:239–246.

Vidal S, Curran J, Kolakofsky D. 1990b. A stuttering model for paramyxovirus P mRNA editing. *EMBO J 9*:2017–2022.

Volchkov VE, Becker S, Volchkova VA, Ternovoj VA, Kotov AN, Netesov SV, Klenk HD. 1996. GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases. *Virology 214*:421–430.

Weiss RB, Dunn DM, Atkins JF, Gesteland RF. 1990. Ribosomal frameshifting from −2 to +50 nucleotides. *Prog Nucleic Acids Res Mol Biol 30*:159–183.

Wu HN, Lin YJ, Lin FP, Makino S, Chang MF, Lai MMC. 1989. Human hepatitis delta virus RNA subfragments contain an autocleavage activity. *Proc Natl Acad Sci USA 86*:1831–1835.