Two nucleotides immediately upstream of the essential A6G3 slippery sequence modulate the pattern of G insertions during Sendai virus mRNA editing

HAUSMANN, Stéphane, et al.

Abstract

Editing of paramyxovirus P gene mRNAs occurs cotranscriptionally and functions to fuse an alternate downstream open reading frame to the N-terminal half of the P protein. G residues are inserted into a short G run contained within a larger purine run (AnGn) in this process, by a mechanism whereby the transcribing polymerase stutters (i.e., reads the same template cytosine more than once). Although Sendai virus (SeV) and bovine parainfluenza virus type 3 (bPIV3) are closely related, the G insertions in their P mRNAs are distributed differently. SeV predominantly inserts a single G residue within the G run of the sequence 5’ AACAAAAAAGGG, whereas bPIV3 inserts one to six G’s at roughly equal frequency within the sequence 5’ AUUAAAAAGGGG (differences are underlined). We have examined how the cis-acting editing sequence determines the number of G’s inserted, both in a transfected cell system using minigenome analogues and by generating recombinant viruses. We found that the presence of four rather than three G’s in the purine run did not affect the distribution of G insertions. However, when the underlined AC of the SeV […]


PMID : 9847338

Available at: http://archive-ouverte.unige.ch/unige:38125

Disclaimer: layout of this document may differ from the published version.
Two Nucleotides Immediately Upstream of the Essential A₆G₃ Slippery Sequence Modulate the Pattern of G Insertions during Sendai Virus mRNA Editing

STÉPHANE HAUSMANN, DOMINIQUE GARCIN, ANNE-SOPHIE MOREL, AND DANIEL KOLAKOFSKY*
Department of Genetics and Microbiology, University of Geneva
School of Medicine, CH1211 Geneva, Switzerland

Received 11 May 1998/Accepted 25 September 1998

Editing of paramyxovirus P gene mRNAs occurs cotranscriptionally and functions to fuse an alternate downstream open reading frame to the N-terminal half of the P protein. G residues are inserted into a short G run contained within a larger purine run (A₃G₅) in this process, by a mechanism whereby the transcribing polymerase stutters (i.e., reads the same template cytosine more than once). Although Sendai virus (SeV) and bovine parainfluenza virus type 3 (bPIV3) are closely related, the G insertions in their P mRNAs are distributed differently. SeV predominantly inserts a single G residue within the G run of the sequence 5′ AAAAAA AAAAAA (whereas bPIV3 inserts one to six G's at roughly equal frequency within the sequence 5′ AUUUUA AUUUUA AUUUUA (differences are underlined). We have examined how the cis-acting editing sequence determines the number of G's inserted, both in a transfected cell system using minigenome analogues and by generating recombinant viruses. We found that the presence of four rather than three G's in the purine run did not affect the distribution of G insertions. However, when the underlined AC of the SeV sequence was replaced by the UU found in bPIV3, the editing phenotype from both the minigenome and the recombinant virus resembled that found in natural bPIV3 infections (i.e., a significant fraction of the mRNAs contained two to six G insertions). The two nucleotides located just upstream of the polypurine tract are thus key determinants of the editing phenotype of these viruses. Moreover, the minimum number of A residues that will promote SeV editing phenotype is six but can be reduced to five when the upstream AC is replaced by UU. A model for how the upstream dinucleotide controls the insertion phenotype is presented.

Sendai virus (SeV), a prototype paramyxovirus, contains a nonsegmented negative-strand RNA genome. This genome RNA of 15,384 nucleotides (nt) is found in a helical nucleocapsid core assembled with a predicted 2,564 copies of the N protein (i.e., if there is but one N subunit per 6 nt, the rule of six), to which 200 copies of the P (phosphoprotein) and 50 copies of the L (large) protein are bound. Paramyxovirus genomes serve as templates for both monocistronic mRNAs and a full-length complementary antigenome strand, the intermediate in genome replication. The nucleocapsid core can carry out mRNA synthesis in vitro, and when it is provided with unassembled N protein, genome replication also takes place. The synthesis of both genomes and antigenomes is coupled to their assembly, and they are consequently found only as nucleocapsids assembled with N protein (reviewed in reference 22).

Six mRNAs (in the order N, P, M, F, HN, and L) are transcribed from the N RNA genome by the P-L polymerase. All of these viral mRNAs except the P-gene mRNA express a single primary translation product from a single open reading frame (ORF). The paramyxovirus P gene mRNAs, in contrast, generally contain alternate ORFs that overlap the N terminus as well as the middle region of the P-protein ORF, and they express several proteins. For SeV, the C-protein ORF overlaps the N-terminal region of the P ORF and is accessed via ribosomal choice during translational initiation (b) (Fig. 1a). 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, i.e., cotranscriptional mRNA editing (3, 29–31).

The subfamily Paramyxovirinae is organized in three genera: respiroviruses (formerly the Paramyxovirus genus), including SeV and bovine parainfluenza virus type 3 (bPIV3); morbilliviruses (e.g., measles and distemper viruses); and rubulaviruses (e.g., mumps virus and simian virus 5). Most of these viral P genes contain an A₃G₅ purine run at the start of the internal, overlapping V ORF (Fig. 1b). mRNAs with expanded G runs are transcribed from these genes in addition to those which are faithful copies of their templates, and the number of G insertions which occur for each virus group mirrors their requirements to switch between the in-frame and out-of-frame ORFs (reviewed in reference 19). For the morbilliviruses and SeV, which require a +1 frameshift to access the V ORF from the genome-encoded P ORF, a single G is added as the predominant insertional event (Fig. 1a). For the rubulaviruses, which require a +2 frameshift to access the remainder of the P ORF from the genome-encoded V ORF, two G's are added at high frequency when insertions occur. For bPIV3, where both V and another ORF (called D) overlap the middle of the genome-encoded P ORF, one to six G's 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 and by 5’ and 3’ end formation (1, 4).

Paramyxovirus mRNAs are made in the cytoplasm, and these viruses consequently must fend for themselves in all aspects of mRNA synthesis. All negative-strand virus RNA polymerases (RNAPs) which polyadenylate their mRNAs are thought to do so by stuttering on a short run of template U residues (4 to 7 nt long), and it was this observation that first suggested that the G insertions would similarly occur by pseudo-templated transcription (3, 17, 29). Paramyxovirus mRNA editing is thought to take place as follows (see Fig. 7a): (i) the viral polymerase is postulated to pause before the end of the template C run (nt 1051 to 1053); (ii) the nascent chain, whose 3’ end is base paired to the template, slips backward by one (SeV and morbilliviruses) or two (rubulaviruses) template positions (Fig. 1); and (iii) as a result, one or two of the template C residues are copied a second time when transcription resumes processively. In the realignment of nascent mRNA and template, U:G (but not A:C) pairs are permitted, and in analogy to ribosomal frameshifting, the region where alternate base pairing occurs after realignment is called the slippery sequence (2, 16, 33). For most viruses, this cycle of slippage and pseudo-templated transcription 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 these different patterns of G insertions (26).

The replacement of the SeV editing region with that of bPIV3 in a SeV minigenome leads to mRNAs with G insertions (26). For each virus group, the upper line shows the mRNA which is an exact copy of the gene, and the beginning of the ORF box indicates the ribosomal start codon. When more than one ORF box is attached to the line, the ORFs are accessed by alternate initiation codons. The boxes below indicate alternate downstream ORFs which are fused into place by G insertions in the mRNAs. The positions of the insertions are shown by the dotted vertical lines. The three possible ORFs are indicated by the dotted brackets. Note that the A run which precedes the G run is the only part of the cis-acting sequence which is strictly conserved according to genera. Also note that the second A residue upstream of the rubulavirus G run is replaced by a G (highlighted with rectangle), which presumably accounts for why rubulaviruses insert a minimum of two G residues when stuttering begins (19, 31). The shaded boxes indicate sequence conservations. When the highlighted AC is changed to UU in the other respiroviruses, SeV edits its mRNA with multiple G insertions, like PIV3.

from that of its template, other than by splicing and by 5’ and 3’ end formation (1, 4).

Paramyxovirus mRNAs are made in the cytoplasm, and these viruses consequently must fend for themselves in all aspects of mRNA synthesis. All negative-strand virus RNA polymerases (RNAPs) which polyadenylate their mRNAs are thought to do so by stuttering on a short run of template U residues (4 to 7 nt long), and it was this observation that first suggested that the G insertions would similarly occur by pseudo-templated transcription (3, 17, 29). Paramyxovirus mRNA editing is thought to take place as follows (see Fig. 7a): (i) the viral polymerase is postulated to pause before the end of the template C run (nt 1051 to 1053); (ii) the nascent chain, whose 3’ end is base paired to the template, slips backward by one (SeV and morbilliviruses) or two (rubulaviruses) template positions (Fig. 1); and (iii) as a result, one or two of the template C residues are copied a second time when transcription resumes processively. In the realignment of nascent mRNA and template, U:G (but not A:C) pairs are permitted, and in analogy to ribosomal frameshifting, the region where alternate base pairing occurs after realignment is called the slippery sequence (2, 16, 33). For most viruses, this cycle of slippage and pseudo-templated transcription 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 these different patterns of G insertions (26).

The replacement of the SeV editing region with that of bPIV3 in a SeV minigenome leads to mRNAs with G insertions whose distribution resembles those found in bPIV3 infections (18). The counting mechanism is thus apparently controlled in large part by a cis-acting sequence. Here we report experiments which show that these controlling sequences lie immediately upstream of the slippery A6G3 purine run and propose a model to account for the counting mechanism.

MATERIALS AND METHODS

Construction of minigenomes. Construction of the internal deletion SeV minigenome (SH22) is described in reference 15. Briefly, pHSH2 is based on pSp65, with the minigenome inserted directly downstream of the T7 promoter and carrying the hepatitis delta virus genomic ribozyme directly after the mini-genome. Each minigenome contains 423 nt of the 5’ trailer/L-gene region of SeV, a 54 nt polynucleoty region into which the 103 nt SeV editing cassette has been inserted between EcoRI and XbaI sites, and 146 nt of the 3’ end of the SeV negative-strand genome including the leader/N-gene region.

The A6G1 series was constructed by inserting the respective P cassettes in place of the corresponding Xba-SeV-EcoRI fragment of SH22 derivatives. To maintain hexamer length, 8 to 13 nt were inserted into the polynucleoty region (15). The A6G1 series were obtained by PCR amplification from pGEM-P HA with primers A6G1 (5’ GACTCTAGAGAGCGACTCAACAAAAAAAGCAT AGGAAGA), A6G2, A6G4, and A6G5 (identical to A6G1 except for the number of G’s at the underlined position), and PECOp (5’ GGGCAGCTTGTGGCAACAC AC). The PCR products were then digested with Xho and EcoRI and introduced in the corresponding SH22 derivatives.

The Swap series was constructed as described above, using for PCR the primers Swap1 (5’ GACTCTAGAGAGGAGACTGGTAAAAAAAGGG), Swap2 (5’ GACTCTAGAGAGCGACTATGGAAATAAAAAAGGG), and Swap3 (5’ GACCT TAGAGAGCGACCTATTAAAAGG). Minigenome expression and direct limited primer extension of the minigenomes. The various minigenomes and the SeV N, P, and L genes were expressed in A549 cells basically as described elsewhere (6, 18). The cells were grown as monolayers in 9-cm-diameter dishes and infected at 2 to 5 PFU/cell with a vaccinia virus expressing T7 RNA polymerase (vTF7-3 [11]). At 1 h post-infection (hpi), the medium was replaced with a transfection mix composed of 20 μl of home-made transfectACE, pGEM-L (1 μg), pGEM-bPIV1-N (2.5 μg), pGEM-bPIV2-N (2.5 μg), pGEM-min-gene (5 μg), and minimal essential medium to 1 ml. After 2 h at 33°C, an extra 6 ml of minimal essential medium was 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–10% Nonidet P-40. Nuclei were removed by pelleting at 12,000 × g for 5 min. To recover mRNA and viral nucleocapsids (6), the cytoplasmic extracts were centrifuged in a step gradient composed of a 5.7 M...
FIG. 2. Minigenome mRNA editing in transfected cells. (a) SeV minigenome replication and transcription was reconstituted in vaccinia virus vTF7-infected cells via the transfection of pGEM plasmids which express the PIV1 N protein (sphere), the P protein with a 10-amino-acid deletion including the editing site (square), and the L protein (oval), as well as the minigenome (see text). T7 refers to T7 RNA polymerase expressed from vTF7-3. The T7 minigenome transcript is assembled with N protein and is replicated and transcribed by the SeV and hPIV1 proteins. The ensuing mRNAs (bottom line) are examined for G insertions by limited primer extension. (b) Limited primer extension analysis of mRNAs and antigensomes from cells transfected with minigenomes containing A5G or editing sites (indicated above the lanes). Cytoplasmic extracts of the transfected cells were prepared at 24 hpi, and their SeV RNAs were separated on CsCl density gradients into fractions containing the mRNAs (pellet) and genomes/antigenomes (banded material) (Materials and Methods). A 5'-32P-labeled primer was then extended across the editing site with RT in the presence of ddATP to limit the extension (schematized above). The intense lower band in each mRNA lane indicates the uninserted mRNAs. The fraction of one-G-insertion mRNAs was determined by densitometry and is shown below the mRNA panel. The results of a parallel transfection of the A5G construct in which pGEM-L was withheld is shown on the left, and the absence of bands here indicates that the other signals observed are dependent on the SeV polymerase. An editing-inactive construct, A4GAG3, was included as another negative control.

RESULTS

Minigenome mRNA editing in transfected cells. (i) The minimum G run is three. We previously used a synthetic minigenome in a cell transfection assay to examine the role of cis-acting sequences on the editing phenotype (18). The viral functions required to replicate and transcribe the minigenome (N, P, and L) are expressed from T7-promoted pGEM plasmids in this system, and the T7 RNA polymerase is provided by coinfection with a recombinant vaccinia virus (vTF7 [Fig. 2a]). The minigenome T7 transcript, a negative strand with exact viral ends, is first assembled with viral N protein, and these negative-strand nucleocapsids are transcribed and replicated by the SeV N, P, and L proteins (Fig. 2a). mRNA (CsCl pellet RNA) is then prepared from these cells, and the presence of G insertions is determined by limited primer extension. As mentioned above, when the minigenomes were engineered to con-
tain the mRNA editing region of either bPIV3 or SeV, the resulting mRNA was found to be edited accordingly (18).

The above system, however, required selective RT-PCR amplification to distinguish the pGEM-P mRNA from the minigenome mRNA. Further, the precise fraction of edited minigenome mRNAs was sometimes unclear due to a background of unedited T7 transcripts present in the minus-pGEM-L negative control. These transcripts presumably arose via recombination between the pGEM-N and the minigenome that contains the first 90 nt of the N gene. To eliminate these problems, pGEM-HAP (a tagged version of P which eliminates C-protein expression) was replaced by pGEM-HAPΔ30, containing a 30-nt deletion around the editing site. This deletion reduces the activity of P ~2-fold (unpublished data), but it eliminates the binding site for the primer used for the limited extension. This primer thus extends only on minigenome transcripts. Second, pGEM-N<sub>hPIV1</sub> was used instead of pGEM-N<sub>SeV</sub> to limit recombination between the pGEM-N and the minigenome. The sequence of the N gene of human parainfluenza virus type 1 (hPIV1) is sufficiently similar to that of SeV to be active in this system, but it is sufficiently different in the first 90 nt to severely limit recombination.

SeV (5′ A<sub>6</sub>G<sub>3</sub>) and b/hPIV3 (5′ A<sub>6</sub>G<sub>4</sub>/5) have slightly different numbers of G's at their editing sites (as for protein-encoding DNA, plus strands are written 5′ to 3′ and minus strands are written 3′ to 5′). We therefore examined the effect of varying the number of G's in this run from one to five. The results of limited primer extension directly on the mRNAs (CsCl pellet) and antigenomes (CsCl band) which accumulated in these transfections of our modified system are shown in Fig. 2b. Two negative controls were included: (i) pGEM-L was withheld from some of the transfections to ensure that all of the signal observed was SeV specific, and (ii) a minigenome with 5′ AAAAGAGGG rather than AAAAAAGGG, known to be inactive in mRNA editing (18) was examined to control for spurious bands. In all cases, the constructs were adjusted to generate minigenomes of hexamer length (by compensating at an EcoRV site ca. 100 nt downstream of the editing site), as otherwise genomes are readily readjusted to hexamer length in this system during antigenome synthesis (genome length correction [15]). Figure 2b shows the relative importance of the exact number of G's in the editing sequence in this modified system. mRNA from a minigenome containing the wild-type (wt) SeV 5′ A<sub>6</sub>G<sub>3</sub> editing sequence contained a single G insertion at 15% frequency (rather than ca. 30% in a natural viral infection [see below]). Its expansion to A<sub>6</sub>G<sub>4</sub> slightly decreased the insertion frequency (to 10%), and this frequency was slightly less than 10% when expanded to A<sub>6</sub>G<sub>5</sub>. In contrast, no edited mRNAs were detected in the A<sub>5</sub>G<sub>1</sub> and A<sub>5</sub>G<sub>2</sub> constructs. In all cases, examination of the CsCl-banded antigenomes showed that no insertions had occurred here (Fig. 2b); hence, the single base insertion found in the minigenome mRNA had occurred during transcription. Efficient G insertion thus occurs when a minimum of three G's are found at the editing site. The presence of four or five G's here is relatively well tolerated and does not lead to a greater fraction of the mRNAs with >1 G insertions.

(ii) The 2 nt upstream of the purine run are a key determinant of the editing phenotype. When 18 nt upstream of the SeV A<sub>6</sub>G<sub>5</sub> run were replaced with the corresponding sequence of bPIV3 and the G run was increased from three to four (as in bPIV3), a significant fraction of the resulting mRNAs contained multiple G insertions, i.e., displayed a PIV3 phenotype (18). We therefore constructed minigenomes Swap2, -5, and -8, in which 2, 5, and 8 nt upstream of the A<sub>6</sub>G<sub>5</sub> run were replaced with the corresponding sequence of bPIV3, to more precisely define the cis-acting sequence responsible for this altered phenotype (Fig. 3). When as few as 2 nt of bPIV3 were placed upstream of the purine run, the minigenome mRNA had a G insertion pattern reminiscent of that found in natural bPIV3 infections. Bands corresponding to two to five G insertions were now detected, and the overall fraction of edited mRNA rose from 15 to >40% (Fig. 3 and 4b). The altered editing phenotype was not influenced by the number of G's in the purine run, as Swap8 constructed in the A<sub>6</sub>G<sub>3</sub> background behaved similarly to Swap8 in the A<sub>6</sub>G<sub>3</sub> background (Fig. 3) as did Swap5 and Swap2 (data not shown). To control that all minigenomes were of hexamer length and that the G insertions had occurred during mRNA synthesis, the CsCl-banded antigenomes were also directly examined. Only a single strong band at the nonedited position was detected (Fig. 3). Thus, the altered insertion pattern of the Swap constructs was due not to the extra G in the polypurine run (5′ A<sub>5</sub>G<sub>4</sub> versus A<sub>5</sub>G<sub>3</sub>) but rather to the differences in the upstream sequence.

mRNA editing in rSeV infections. The vTF7-infected/plasmid-transfected cell system, although highly artificial, has the advantage that the translational consequences of the mRNA editing do not feed back on the editing process itself, as all viral proteins (N, P, and L) are provided via T7 RNA polymerase. mRNA editing can thus be studied in isolation. However, unlike the case for natural infections, the transfected cells constitutively produce large amounts of viral proteins whose stoichiometric balance is not subject to viral regulation, and this may also affect editing. Moreover, in our modified transfection system, the N<sub>hPIV1</sub> and HAPΔ30 genes are used in place of N<sub>SeV</sub> and HAP, which reduces the editing frequency. To study mRNA editing in natural virus infections, we constructed rSeV containing mutations at the editing site (Materials and Methods). The amino acids of the P sequence altered by the editing mutations (Fig. 5) appear to lie in a noncritical region of the
protein (7), as all of these rSeV were prepared as readily as the wt control. A549 cells infected with these viruses all accumulated similar amounts of N and M protein (at 24 hpi) as judged by immunoblotting, and similar levels of genomes and antigens as judged by primer extension, as the wt control (data not shown).

Total RNA was isolated from cells infected with these editing mutants as well as SeV, bPIV3, and an rSeV, and the pattern of mRNA editing was examined by limited primer extension. The primer was also extended on SeV DNA and Swap8 DNA (used to prepare the viruses) as negative controls. The results obtained with natural virus infections of rSeV-Swap2, -5, and -8 were basically similar to those obtained with the minigenomes in transfected cells (Fig. 4). In contrast to the SeV and rSeV infections, where the insertion was mostly limited to a single G and mRNAs with >1 G insertions represented only 1.5% of the population, the substitution of only 2 nt was sufficient to lead to multiple G insertions in 36% of the mRNAs (Fig. 6b). The further substitution of 5 and 8 nt led to a more even distribution of 1 to 15 insertions, as in the natural bPIV3 infection. Not any change within the upstream 8 nt led to a PIV3-like phenotype; e.g., mutating the 8 upstream nt to 5′ UCACCCUG (mostly the complement of the bPIV3 sequence) did not lead to detectable >1-G insertions (data not shown). All the rSeV-swap constructs edited a greater fraction of the mRNA than did either SeV or bPIV3, and the frequencies of one-G insertions are clearly increased here as well. Thus, the insertion of multiple guanylates at the editing site, the hallmark of PIV3 phenotype, appears to be governed by the upstream sequence. The two bases directly upstream of the 5′ A₆G₃ purine run (UU for h/bPIV3 versus AC for SeV [Fig. 1b]) appear to be key elements of this cis-acting sequence. We also note that neither the minigenome nor rSeV-Swap8 pattern precisely mimics that of the natural bPIV3 infection (Fig. 4b).

Alteration of the A run and mRNA editing. Morbilliviruses, which, like SeV, edit their P mRNAs by the insertion of predominantly a single G residue, are a more closely related group of viruses than the respiroviruses. Morbilliviruses also contain a more strongly conserved sequence at the editing site (5′ YCC AUU A 5′ G3) than the respiroviruses, where only 5′ A₆G₃ is

![Image](image-url)
conserved (Fig. 1b). As morbilliviruses contain a run of only five adenosines, we were interested in the editing phenotype of a SeV which contained the morbillivirus cis-acting sequence. This was accomplished by converting the sequence 5′ A6G3C to A5G3CC (Fig. 6) so that (i) hexamer genome length is maintained and (ii) the hexamer phase of the first G of the G run moves upstream by one position, which is in fact to that conserved in morbilliviruses (20). The concomitant amino acid substitutions in the P protein are shown in Fig. 5. Again, these viruses were prepared readily and grew to similar levels in eggs as wt virus (data not shown).

Total RNA was isolated from cells infected with each rSeV and examined by limited primer extension (Fig. 6). Shortening the A run from six to five within the context of the SeV upstream sequence (5′ AAC A5G3) eliminated G insertion. When the upstream two bases are further changed to those of the morbilliviruses (5′ AUU A5G3), editing activity is regained. However, the mRNAs are now edited in a pattern similar to that of bPIV3 rather than that of SeV or the morbilliviruses (11). These results are found at roughly equal frequency and together represent 44% of the mRNAs). The substitution of the single upstream nucleotide within the A5G3 background (AAU A5G3) was also sufficient to restore some editing activity, but the overall fraction of inserted mRNAs (24%), as well as the extent of the insertions (>1 G, 11%), was reduced. The SeV polymerase can indeed stutter with an A5 run, but only when U rather than C precedes this run. Efficient stuttering, however, requires the replacement of both upstream nucleotides. Thus, shortening the A run from six to five in SeV consistently decreases the fraction of edited mRNAs as well as the number of G insertions per molecule of mRNA. On the other hand, within the A5G3 background, substitution of U for the upstream C has the opposite effect, and the substitution of UU for the upstream AC has an even stronger effect in promoting G insertions during mRNA synthesis. Remarkably, SeV carrying the morbillivirus cis-acting sequence (AUU A5G3) was found to edit its mRNA in a manner similar to that of bPIV3 rather than that of the morbilliviruses.

**DISCUSSION**

An interesting feature of paramyxovirus mRNA editing is that each particular virus inserts G residues in a pattern matched to the organization of its P gene ORFs (Fig. 1a). Furthermore, the SeV transcription complex can be induced to switch its editing phenotype, by simply substituting the bPIV3 editing region for that of SeV. Our results have highlighted the importance of the sequence immediately upstream of the A5G3 slippery sequence for this phenotype switch, rather than the number of G's in the G run. A minimum of three G's were found to be required for SeV editing to occur, and increasing the G run to four or five was tolerated and did not affect the editing phenotype. We have not examined the effects of further extending the G run. However, a recombinant measles virus engineered to contain three extra G's at the editing site (5′ AUU A5G6) was found to insert two to four G's with increased frequency (28).

One unexpected finding of this work is that replacing the SeV slippery sequence (A5G3) with that of measles virus (A5G3) inactivates the editing process, unlike the analogous h/bPIV3 (A5G4–5) swap. It is possible that SeV stuttering is adversely affected by a run of only five adenylates and/or by the displacement of the hexamer phase of the G run with respect to the N-protein subunit upstream by 1 position (to shorten the A run) (20). However, neither of these possible requirements appears to be critical for stuttering, as replacement of the upstream AC with the conserved UU restores the G insertions.
FIG. 7. A stuttering model for paramyxovirus mRNA editing. (a) The putative RNA-RNA hybrid between the polyprimidine tract of the negative-strand genome (top strand, written 3′ to 5′) and the polypurine run of the nascent mRNA chain (bottom strand, written 5′ to 3′) when the active site of the transcription complex has just incorporated G1052 (top left). At the editing site (shown as C1052, highlighted with a gray box), the transcription complex has the choice of realigning its nascent RNA chain upstream on the template before the next nucleotide is added. Here, the minimum requirement for stuttering, that the realigned hybrid be nearly as stable as its predecessor, has been met because non-Watson-Crick U:G pairs (highlighted in gray) do not disrupt the helical stack. If the rate constant for pseudo-templated G insertion (shown as C1052 in Fig. 7). This template C is then copied a second time when nucleotide addition recurs, resulting in a pseudo-templated G insertion.

Most unexpectedly, significant amounts of mRNAs with multiple G insertions were generated during this restored editing. This unexpected switch to the PIV3 phenotype highlights how little we know of how the cis-acting sequence determines the distribution of G insertions. It also suggests that there may well be differences in how the various viral polymerases interact with the cis-acting sequence and/or that we have as yet identified only one element of the cis-acting sequence, that which has been conserved. Despite our limited information, the decision of the viral polymerase to stutter or not to stutter, and the extent of the stuttering, can be considered in terms of a competitive kinetic model, as this does not depend on detailed structural information of the transcription elongation complex.

A competitive kinetic model for paramyxovirus polymerase stuttering. Cellular RNAPs respond to intrinsic signals in the template DNA and nascent RNA which divert a fraction of the transcription complexes from the path of rapid chain elongation, e.g., to pause or to terminate the chain (23). These processes are among the best-studied examples of transcriptional processes and nascent chain release (1 to 10 s), and the barriers to elongation and termination are roughly equal (16 kcal/mol). The barrier to termination at these positions, however, is thought to be >30 kcal/mol (32). Transcription thus occurs with an infinitesimal probability of spontaneous termination at nonterminator positions. At terminator positions, however, the length of time the polymerase pauses at this site is roughly equal to that for nascent chain release (1 to 10 s), and the barriers to elongation and termination are roughly equal (16 to 18 kcal/mol). By analogy, the barrier to realignment at nonstuttering sites is expected to be very high because of the instability of the realigned hybrid at heteropolymeric sequences, and stuttering does not occur.
We assume that the barrier to nascent chain realignment at stuttering sites is strongly reduced (and lower than that to strictly templated elongation), such that a significant fraction of the elongation complexes are realigned upstream on the template before the next nucleotide (G1053 [Fig. 7]) can be added.

The competitive kinetic model made two predictions for the E. coli elongation/termination decision (32), which should apply as well to stuttering. First, stuttering should be possible only at sites where the realignment complex is relatively stable, such that the relative heights of the elongation and realignment barriers are now reversed. Because of the large difference in barrier heights to realignment at (heteropolymeric) nonstuttering positions, the position of stuttering sites should be strongly determined and the elongation-stuttering decision should have the character of a binary switch. This is of interest because paramyxovirus polymerase stuttering (mRNA editing or polyadenylation) normally does not occur during antigenome synthesis, where the viral polymerase is said to be switched to the replication mode by the concurrent encapsidation of the nascent chain. Second, editing efficiencies should be easily modified (or regulated) at editing sites, because relatively small changes in either kinetic or thermodynamic components of the activation energy barriers (e.g., by the UU/AC substitution) will produce large changes in editing efficiency due to the exponential form of the relationship (20a, 32).

For SeV, there is evidence that the editing site is the middle C (nt 1052) of the template 3' UUCG pyrimidine run (reference 31 and unpublished data), and we assume 7 bp between the nascent chain and the template (Fig. 7); the E. coli complex is thought to contain 8 bp (23–25). The SeV transcription complex is proposed to pause at the editing site, and the polymerase reaction center and nascent RNA 3' end are proposed to realign upstream on the template by one position before the incorporation of G1053. The new 7-bp hybrid formed on realignment is only slightly less stable than its predecessor, as one G:C pair has been replaced by a G:U pair (a difference in stability of ca. +1 to +2 kcal/mol, ignoring the protein components). If the barrier to realignment is lower than that to elongation, the reaction center will equilibrate between these two positions (nt 1051 and 1052) according to the relative stability of the hybrids (Fig. 7). If the realigned hybrid is 1 kcal/mol less stable than the original hybrid, the active site would be found ca. 20% of the time at the realigned position (nt 1051) and 80% at the strictly templated elongation position (nt 1052). If the barrier heights to normal and pseudo-templated elongation are equal, a single G will then be inserted into the mRNA with a frequency of ca. 20%. Of course, after one round of stuttering, the transcription complex is now back to where it started. If nothing else has changed, the process will be repeated at the same frequency. This situation best describes wild-type virus (5' AACC G3) mRNA editing, where insertion of a single G is by far the predominant editing event.

The cis-acting editing sequence appears to control two distinct branchpoints (Fig. 7). In the first, the transcription complex arrives at a site that allows realignment of the nascent chain upstream, and this decelerates the rate of strictly templated nucleotide addition (kforward). If an NMP is nevertheless added, the polymerase can move past the potential site of editing (to escape) (Fig. 7). This first branchpoint determines the overall efficiency of editing. However, if a stutter occurs, the frequency with which it now escapes to strictly templated elongation or restutters represents a second branchpoint, because h/bPIV3 (and SeV-AUUA A6G3) behave differently at this juncture. The second branchpoint determines the number of G insertions once stuttering has commenced. There is experimental evidence that kstutter and krestutter represent separate parts of this reaction pathway. Partial substitution of the GTP in in vitro reactions with ITP (expected to destabilize the nascent chain:template G:C pairs at the editing site), and very low concentrations of GTP (expected to increase the step time for elongation at the editing site), increases the fraction of mRNAs with one G inserted ca. 2-fold (from 20 to 40%), but increases that with multiple G insertions ca. 10-fold (from 2 to 20%) (31). Our results provide further evidence that the second branchpoint can be modulated, during natural SeV infection, by substituting UU for the AC immediately upstream of the 5' A6G3 site. According to a competitive kinetic model, this substitution would presumably lead to a more stable (rather than a less stable) realignment complex, as the upstream UU would somehow more than compensate for the additional U:G pair in the hybrid on realignment. The reaction center would thus now be located mostly at the realigned position (lower ball at position 1051 in Fig. 7), ensuring that stuttering would occur repetitively, and that escape to strictly templated transcription (incorporation of G1053) would be infrequent.

For the E. coli RNAP transcription complex, the nascent chain becomes available for annealing with short oligonucleotides when it is 14 to 16 nt from the RNA 3' end (21, 27). The 6 to 8 nt upstream of the 8-bp RNA-DNA hybrid are thought to be located in an exit channel (or tunnel), also referred to as the tight product-binding site (5, 25). Alteration of the RNA-protein interactions in the E. coli exit channel due to the formation of hairpin structures are thought to somehow stabilize the paused conformation (5a). Similarly, the nascent RNA of the vaccinia virus RNAP elongation complex exits from the polymerase 16 to 18 nt from the RNA 3' end (10, 14). This RNAP was recently found to stutter at the end of a template A9 run (if the templated guanylate immediately following the run was disfavored by severely limiting the GTP concentration), adding one to seven pseudo-templated uridylates to the RNA (9). These slipped RNAs were released from the template, and Deng and Shuman (9) have speculated that the RNA-protein interactions that normally stabilize the nascent chain are weakened when poly(U) occupies most of the RNA exit channel on the vaccinia virus polymerase. Assuming that the basic structural features of all RNAPs have been conserved throughout evolution, the eight bases upstream of the purine run would also be included in the SeV polymerase exit channel. The additional stability conferred by the upstream UU can then be postulated to result from base-specific interactions of the SeV RNAP and the nascent chain as the latter traverses the elongation complex. As long as these base-specific interactions with the exit channel can occur, repetitive G insertions are favored. Each additional G insertion, however, moves the upstream UU toward the outside of the exit channel. These base-specific interactions (and the additional stability of the realigned hybrid) will eventually no longer be possible. The reaction center will then revert to spending most of its time at nt 1052 rather than nt 1051, and the polymerase will soon escape to strictly templated transcription. The counting mechanism by which PIV3 inserts one to six G's at roughly equal frequency would then be based on the dimensions of this RNAP exit channel.

ACKNOWLEDGMENTS

We are grateful to Hans Geisselmann (Grenoble, France) and Kyle Tanner (Geneva, Switzerland) for helpful discussions on thermodynamics.

This work was supported by a grant from the Fonds National Suisse.
REFERENCES