Partial Characterization of a Sendai Virus Replication Promoter and the Rule of Six

THIERRY PELET, CHRISTOPHE DELENDA, OLIVER GUBBAY, DOMINIQUE GARCIN, and DANIEL KOLAKOFSKY

Department of Genetics and Microbiology, University of Geneva School of Medicine, CMU, 9 Avenue de Champel, CH1211 Geneva, Switzerland

Received February 26, 1996; accepted July 23, 1996

We have used a cDNA copy of a natural, internally deleted, Sendai virus defective interfering genome to study the effect of insertions and deletions (which maintain the hexamer genome length) on the ability of viral genomes to be amplified in a transfected cell system. The insertion of 18 nt at nt2 (in the 5' untranslated region of the N gene, just downstream of the le region) was found to be lethal, whereas similar insertions further from the genome ends were well tolerated. Curiously, the insertion of 6 nt on either side of the le/N junction (at nt42 and nt67) was well tolerated, but the insertion of 12 nt at either site, or of 6 nt at both sites, largely ablated genome amplification. These results suggest that an element of this replication promoter is located downstream of nt67, in the 5' untranslated region of the first gene. Remarkably, the addition of 6 nt by the insertion of 2, 3, or 4 nt at nt+7 plus the insertion of 4, 3, or 2 nt, respectively, at nt-7 was poorly tolerated, presumably because the hexamer phase of the intervening sequence was altered with respect to the N subunits of the template. These results suggest that the rule of six operates, at least in part, at the level of the initiation of antigenome synthesis.

Nonsegmented negative-strand (–) RNA viruses have recently been grouped together in the Mononegavirales superfamily, as the similarities in the way Rhabdoviridae, Paramyxoviridae, and Filoviridae genomes are organized and expressed have become clearer (Bishop and Pringle, 1995). The genomes of these virus families vary from 11 to 19 kb in length and express from 5 to 10 genes (or of ongoing (N) protein synthesis, presumably because genome synthesis and assembly are coupled. Under these conditions, the pool of unassembled VSV N protein is chased into nucleocapsids (Arnheiter et al., 1985). In contrast, synthesis of the unassembled le RNA and le RNAs, as well as of the mRNAs, continues (or increases) (Blumberg et al., 1981) (cf. Fig. 1). Initiation of N protein assembly on the nascent le RNA (and le RNAs) is thought to control genome replication, as concurrent assembly of the nascent chain is necessary for the viral replicase to read through the le/N (or le/L) junction and for the processivity of the polymerase to be maintained (Vidal et al., 1989; Kolakofsky et al., 1991).

Much of what we know about mononegavirales genome synthesis is due to work with rhabdovirus and paramyxovirus defective interfering (DI) genomes. DI genomes are deletion mutants generated by a copy-choice mechanism, in which the viral replicase (carrying its nascent chain) either jumps forward on its template creating an internal deletion (e.g., E307, Fig. 1), or jumps back on the nascent chain itself, creating a terminal complementary repeat upon finishing the chain (e.g., H4, Fig. 1) (Leppert et al., 1977; Perrault, 1981, Lazzarini et al., 1981; Re, 1991). These latter DIs are called copy-backs, and all natural examples have replaced their le regions (and a variable amount of sequence downstream, cf Fig.1)
with those of the \( \text{le}^- \) region, and therefore contain the \( \text{le}^- \) region at both ends. Copy-backs tend to have a competitive advantage in replication over the nondefective genome (and internal deletion DI genomes such as E307), presumably because the \( \text{le}^- \) region contains a more efficient promoter for genome replication than the \( \text{le}^- \) region. These notions of relative promoter efficiency have recently been given more substance by Calain and Roux (1995) and Tapparel and Roux (1996), who compared the amplification levels of SeV copy-back and internal deletion DI constructs in the vTF7-3-infected/transfected cell system (Fuerst et al., 1986). The relative amplification advantage of the copy-backs was found to be determined primarily by whether the H4 (or E307) DI genome contained the \( \text{le}^- \) region at both ends, independent of whether or not they expressed mRNAs. The \( \text{le}^- \) promoter was also found to be more sensitive to inhibition by the viral C protein(s), which further accentuates the difference in the relative efficiencies of the \( \text{le}^- \) and \( \text{le}^- \) promoters (Cadd et al., 1996).

The template for paramyxovirus RNA synthesis is the N protein (formerly NP protein), RNA nucleocapsid, in which each N subunit is associated with precisely 6 nt (Egelman et al., 1989), and this is probably why the hexamer length of a SeV genome chain is so important for the efficiency of genome amplification ("the rule of six"; Calain and Roux, 1993). Assembly of genome and antigenome chains is thought to begin flush with the conserved sequences at the 5' ends of the nascent \( \text{le}^- \) and \( \text{le}^- \) RNAs. As 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. Presumably the viral polymerase initiates more efficiently when, e.g., the first six bases of the promoter sequence (3'\text{OH}_\text{UGGUUU}) are found within the same subunit. For paramyxoviruses and rhabdoviruses, the promoter for antigenome synthesis presumably includes not only the cis-acting sequences at the 3' end of the (--) genome template (where the RNA initiates), but also their complement at the 5' end of the nascent antigenome chain, where nucleocapsid assembly begins. (We use the term "leader" to refer to both complementary sequences of this region, and \( \text{le}^- \) promoter refers to all the promoter elements, including those that might be found in the adjoining genic regions.) In analogy to HIV RNA synthesis from proviral DNA, where promoter elements (TARs) are found at the 5' end of the nascent RNA chain as well as on the template (Cullen, 1990; Marciniak and Sharp, 1991; Laspia et al., 1993), concurrent assembly of the nascent mononegalovirus RNA is thought to promote antigenome synthesis by ensuring the processivity of the viral replicase (Vidal et al., 1989).

The extents of these cis-acting sequences, of either the RNA synthesis components or the assembly components of the replication promoters, are largely unknown. Only the first 12 conserved and complementary nucleotides at the ends of the genomes/antigenomes of each Paramyxovirinae genus are clearly involved (for SeV, 3'\text{UGGYYU GUUCUC} is found at the 3' ends of both genomes and antigenomes). The remainder of the \( \text{le}^- \) and \( \text{le}^- \) regions are always very AU-rich, but there is no strong conservation of sequence either between the \( \text{le}^- \) and \( \text{le}^- \) regions of each virus or between each region of closely related viruses. Blumberg et al. (1991), however, have pointed out a semiconserved sequence ca. 75-95 nt from the ends of several paramyxovirus genomes/antigenomes (the BB box), which could be involved in RNA synthesis. For SeV, the replication promoters appear to be largely confined within the 5' 120 nt (up to the N protein AUG) and the terminal 145 nt (near the end of the L gene) of the antigenome, as DI analogues carrying a foreign gene (CAT) and containing only these SeV sequences are viable (Park et al., 1991). The numbering convention used starts at the \( \text{le}^- \) region and ends with the \( \text{le}^- \) region of both (+) antigenomes and (--) genomes,
similar to the convention for numbering sequences in dsDNA, and consistent with the direction of gene expression. The extent of the le region promoter can also be gauged by the minimum size of the replacement of the le region with the le region in copy-back DI genomes. By this criterion, only 45 nt of the 46-nt rhabdovirus VSV le region are required for efficient replication (Perrault, 1981; Kolakofsky, 1982), and recent reverse genetic studies have further reduced this minimal sequence to the first 36 nt of the le region (Pattnaik et al., 1995). For paramyxovirus copy-back DI genomes, on the other hand, the minimum replacement has so far been 110 nt for SeV (Leppert et al., 1977; Calain et al., 1992) and 94 nt for measles virus (Sidhu et al., 1994), indicating that this paramyxovirus replication promoter may be more extensive. This paper reports on the effects of insertions and deletions in DI-E307 (which maintain hexamer genome length) on their ability to be amplified. These studies confirm the hypothesis that an element of the le region promoter lies within the adjacent N gene (downstream of nt72) and that another element within nt67-67 is governed by its hexamer phase.

MATERIALS AND METHODS

Cells, transfections, and plasmids

A549 cell monolayers seeded on 5-cm petri dishes were infected with vaccinia virus recombinant expressing T7 RNA polymerase (vTF7-3; Fuerst et al., 1986) at 2–3 PFU/cell. One hour postinfection, the medium was replaced with 2 ml of MEM containing 20 μl of Transfect-Ace (Rose et al., 1991) and the following amounts of pGEM plasmids: 3 μg of N, 3 μg of PHa, 1 μg of L, and 5 μg of pE307 derivatives. All plasmids were described previously (Curran et al., 1991; Engelhorn et al., 1993). Transfected cells were maintained in serum-free medium at 33°C for 36 hr before harvesting.

Construction of mutants

BgIII and Nsil restriction sites were introduced in pE307 (pE307A) by the gapped plasmid method (Curran and Kolakofsky, 1991). Mutants Nsi/Not*, Nco/Not*, Kpn/Not*, Bgl/+12, and Bgl/+18 were generated by inserting the following partially self-complementary (in boldface type) oligonucleotides at these sites: Nsi/Not*, 5’-TAT-GCGGCCCATGATCA 3’; Nco/Not*, 5’-CATGGATGC-GGCGCCATC 3’; Kpn/Not*, 5’-CATGGCGGCCCATGG- TAC 3’; Bgl/+12, 5’-GATCTTCGCGAA 3’. Mutant Bgl/+18 was made by the insertion of the Bgl/+12 oligonucleotide into the +6/0 mutant. The insertion of these oligonucleotides duplicated restriction sites, thereby facilitating the generation of the corresponding revertants (Nsi/ΔNot, Nco/ΔNot, Kpn/ΔNot, Δ12, and Δ18) by digestion with the relevant restriction enzymes, followed by religation.

Mutants 0/+6, +6/+0, +6/+6, +3/+3, +3/+3, +1/+5, and +5/+1 were all made using the polymerase chain reaction (PCR). Depending on the mutant, PCR primers were used with insertions or deletions, which overlapped the BgIII site, the Nsil site, or both sites. The PCR products were then digested with BgIII/Nsil and subcloned into the corresponding sites of pE307A.

Mutants +2/−2, +2/−4, and +4/−2 were made by the “single-oligo” method. Briefly, pE307A was digested with BgIII/Nsil, treated with alkaline phosphatase, and purified from an agarose gel. The resulting linear vector carried a 3’ overhang at one end and a 5’ overhang at the other end. Oligonucleotides spanning the entire BgIII–Nsil region (and containing the insertions/deletions) were synthesized so that their 3’ and 5’ terminal nucleotides were complementary to both the 5’ and the 3’ overhangs of the pE307A vector. Following phosphorylation of the 5’ ends with T4 polynucleotide kinase, 30 pmol of oligonucleotide was annealed to approximately 0.1 pmol of vector (ratio 300:1) for several hours at 16°C. A fill-in/ligation reaction was subsequently performed on the annealing products in 50 mM Tris, pH 7.4, 10 mM MgCl2, 10 mM DTT, 1.5 mM ATP, 0.5 mM dNTPs, 1 unit of Klenow enzyme, and 1.5 units of T4 DNA ligase, at 16°C overnight. Transformed colonies were streaked onto nitrocellulose filters and screened by hybridization with 32P-labeled oligonucleotides, similarly to the method described in Vidal et al. (1989).

Regardless of the methods employed, all the mutations were confirmed by DNA sequencing.

RNA purification, primer extensions, and Northern blotting

Twenty-four to 48 hr posttransfection, cell extracts were prepared and nucleocapsids were isolated on CsCl density gradients (40,000 rpm for 90 min at 12°C in a SW60 rotor). The banded nucleocapsids were diluted in 7 volumes of TE (10 mM Tris–Cl, pH 7.4, 1 mM EDTA) and pelleted in a SW60 rotor at 12°C for 1 hr at 50,000 rpm. The purified nucleocapsids were then resuspended in 0.1% sarkosyl/TE and phenol extracted, and the RNA was recovered by ethanol precipitation.

Primer extensions were performed essentially as described in Garcin et al. (1995). The relative amounts of (−) genomes present in the various CsCl-banded nucleocapsids were determined by extension of primer L15270 (5’-GAAGCTCCGCGGTACC 3’) for Nsi/Not* and Nco/Not*, and primer L15203 (5’-GGATCAGTGGTGATGATTG 3’) for Kpn/Not*. (+) RNAs were quantitated with primer NP126 (5’-CGGCCATCGGCTACAGT 3’). Primer extension products were separated on 6% acrylamide sequencing gels.

For Northern blotting, nucleocapsid RNAs were loaded
RESULTS

Some insertions of hexamer length inactivate the DI-E307A genome

pE307 (Engelhorn et al., 1993) contains a copy of a natural DI genome with a single deletion of 13,590 nt fusing the beginning of the N gene to the end of the L gene (see Fig. 1) [and generating a (+) transcript from its T7 promoter with three extra 5' G residues]. pE307 was modified at four positions creating a BglI site and a NsiI site on either side of the le'+/N junction (E307A, Figs. 2 and 4). These four substitutions did not affect the efficiency with which this genome was amplified from cDNA in the transfected cell system. A full-length DNA based on pE307A (FL-3; Garcin et al., 1995) could be rescued into infectious virus with remarkable efficiency. However, another full-length construct, marked by the NsiI site at nt 72 [5' CCGG of the 12-nt insertion (in capital letters), in the 5' untranslated region of the N gene, just downstream of the NsiI site], could not be rescued into virus (unpublished results). The same insertion, moreover, was found to strongly suppress (or eliminate) the ability of DI-E307 to be amplified in the transfected cell system (lane +12 at 77, Fig. 4). We assumed at first that this inhibitory effect was due to the very GC-rich nature of the insertion, but we were unable to recover activity by removing the central 5' GGCC of the 12-nt insertion (by mung bean nuclease digestion of the BspMI-cut DNA) and their replacement with 5' ATAT at nt 72, to restore hexamer genome length (not shown). The inhibitory effect of the 12-nt NsiI insertion was also unexpected, as a 42-nt polylinker can be placed in the BspMI site of E307 without eliminating its activity in amplification (Mottet et al., 1996).

To further examine this inhibitory effect, 18-nt insertions containing NotI sites were placed at nt 78 [the number refers to the first position whose base has been changed, rather than the cleavage site] (NsiI), nt 852 (NcoI), and nt 1794 (KpnI) of pE307A (1794 nt, Fig. 2). Insertions of 18 nt were used so that, in each case, the particular site could also be duplicated (e.g., 5' TGGCATNNGCGGCCTGAgga was placed in the NsiI site; the insertion is in capital letters) (Materials and Methods). The revertant could thus be easily prepared by restricting and religating the plasmid DNA, to control for inadvertent mutations during construct preparation (especially those which would change the hexamer genome length). The various pE307A constructs were transfected into vTF7-3-infected cells along with pGEM plasmids expressing the N, P, and L proteins. Cytoplasmic extracts were pre-

![Image](https://via.placeholder.com/150)

FIG. 2. The effect of 18-nt insertions on the amplificability of DI-E307A. The E307A antigenome is schematized as a horizontal line (not drawn to scale), with three additional 5' guanosines (in lower case) derived from the T7 promoter. Short vertical lines above the antigenome mark the le'/N mRNA and L mRNA/le' junctions. The ORF of the N/L mRNA is shown as an open box, in which the slanted, dotted line shows the fusion of the two ORFs at nt 607. The relative positions of restriction sites used to introduce insertions are shown below; nucleotide positions are indicated in brackets. The relative positions of the primers used to detect the 5' ends of the (-) genomes and (+) RNAs are indicated with heavy horizontal arrows. Note that the AG + 3 band of the (NsiI/NotI +) construct, and the (-) genome band of the KpnI/NotI+ construct, are at a higher position than the others, because the insertion here is in between the primer and the 5' end of the chains. The various constructs, containing either the 18-nt NotI insertion (NotI+:) at the indicated site, its reverted revertant (NotI:Δ), or the unmodified E307A DI genome (lane wt), were examined for their ability to be amplified in vTF7-3-infected/transfected cells (see text). As a negative control, pGEM-L was not included in some of the infections, as indicated above. The relative amounts of (-) genomes present in the various CsCl-banded nucleocapsids were determined by primer extension (Materials and Methods). The positions of the 5' ends of the T7 polymerase generated antigenome (AG + 3), of the SeV polymerase generated antigenome (AG), and of the pGEM-derived N mRNA (N) are indicated on the right. The arrowhead next to the third lane of the (+) RNAs shows the expected position of the SeV-generated antigenome.
pared from these cells after 24–48 hr of incubation, their nucleocapsids were isolated on CsCl density gradients, and the relative amounts of E307A genomes and anti-genomes present in these fractions were estimated by primer extension. To ensure that the E307A RNAs detected were amplified by the SeV polymerase, pGEM-L was omitted from some of the transfections as a negative control. Analysis of E307A (+) antigenomes and (−) genomes gave similar results, as multiple rounds of replication occur during these transfections, e.g., at least as many antigensomes had accumulated due to replication by the SeV enzymes as with those made by T7 polymerase and assembled with N protein (cf. middle panel of Fig. 2). Analysis of the (−) genome levels in this assay is presented first, as these RNAs are generated only by the SeV polymerase. As shown in Fig. 2 (bottom), E307A genomes could not be detected for the construct with the insertion at nt72 (Nsi/Not+), whereas normal levels (relative to the wt E307A control, lane wt) were found in cells transfected with this revertant (Nsi/ΔNot), and these were dependent on the cotransfection of pGEM-L. Analysis of the levels of (+) RNAs found assembled with N protein is more complex, because the (+) RNAs are generated by both the SeV and the T7 RNA polymerases. The T7-generated antigenome (AG + 3, Fig. 2, middle) can, however, be distinguished from that made by the SeV enzyme (AG), because it contains three extra guanosines at its 5′ end (shown in lowercase letters, Fig. 2) which are eliminated during its replication, presumably by the SeV enzyme. E307A antigenomes could also not be detected for Nsi/Not transfection (its estimated position is indicated by an arrowhead, by reference to a sequence ladder run alongside, but not shown), whereas normal levels were found in cells transfected with Nsi/ΔNot, again dependent on the cotransfection of pGEM-L. There is also a strong 5′ end below the position of the antigenome band, which corresponds to the end of the transcript generated from pGEM-N and which (like the AG + 3 band) is absent from natural virus infections (not shown). The assembly of the pGEM-N mRNA by its translation product is common in our recovery systems and is probably another example of the uncoupled assembly of RNAs in the transfected cell cytoplasm. The levels of these pGEM-N mRNAs serve, along with the AG + 3 bands, as internal controls for judging RNA recovery. The levels of the SeV polymerase amplified antigenomes can be determined by reference to these T7 polymerase transcripts. In contrast to the deleterious effects of inserting 18 nt containing a NotI site at nt47, similar insertions at the Ncol and KpnI sites were seemingly without effect, as in both cases they were found to be amplified as efficiently as their respective revertants (Fig. 2).

The above results suggest that it may be the location of the 18-nt insertion at nt47, i.e., its proximity to the le+ region, that is deleterious to E307A amplification. The 18-nt insertion at nt1695 (KpnI) may be without effect as it is too far away from the other end of the 1794-nt DI RNA. We wished to examine whether insertions within the le+ region itself were similarly deleterious to DI genome amplification. However, as this region is likely to contain sequences important for genome replication, we inserted 6, 12, and 18 nt, and more similar to normal base composition, near the BglII site (Fig. 4). The BglII site was also duplicated by the 12- and 18-nt insertions, so that the revertant could again be prepared as a control, if the insertion proved deleterious. As shown in Figs. 3 and 4, the addition of six bases at nt47 had only a small effect on the amplification of E307A (2-fold in Fig. 3, none in Fig. 4), whereas the insertion of 12 nt strongly suppressed its amplification (>30-fold), and the insertion of 18 nt almost completely eliminated its activity. These latter deleterious effects, moreover, were not due to inadvertant mutations elsewhere, as the subsequent elimination of the insertions fully restored the amplification levels in both cases (Fig. 3).

Evidence for a replication promoter element within the N gene

The 12- and 18-nt insertions at nt47 were chosen with an eye to minimizing the possibilities of RNA secondary structure, yet they were poorly tolerated. The further ability of E307A to tolerate the insertion of 6 nt, but not 12 or 18 nt at nt47, suggests that it is not the sequence here
be tolerated, but its displacement by 12 or 18 nt is not. The ability of E307 to tolerate a 42-nt insertion at nt72 remains unexplained, but could be due to the partial duplication of nt73-96, which contains the common sequence 5’ 73ctgAR GNYRR GYTCYN GANCYN96 (spac ing denotes hexamer phase). For comparison, the conserved sequence found at 77–96 nt from the 5’ ends of genomes and antigenomes (the BB box) can be similarly written as 77AR GRRCA GTYCN GACYYY96. Moreover, the possibility that a 42-nt insertion (but not a 12- or 18-nt insertion) can be tolerated at this position, independent of the sequence, cannot be excluded.

The le+ replication promoter and the rule of six

We also modified E307A by the insertion of 2 nt at nt47, and the compensatory deletion of 2 nt at nt67 (+2/−2) or the insertion of 4 nt at nt67 (+2/+4) (Fig. 5), so that the DI genome remained of hexamer length. E307A−2/0 was found to be amplified well, whereas E307A−2/4 was amplified very poorly. We at first assumed that E307A−2/4 contained an inadvertant nonhexamer deletion or insertion elsewhere which was responsible for its poor replication activity. However, the replacement of nt47-67 of

which is important for genome replication (the sequence of the BglII site also contains two base changes from the wt sequence, Fig. 4). These results are more consistent with the length of the insertion per se in this region being important. We therefore examined whether the insertion of 6 nt just upstream of the Nsi site would be tolerated, and the effect of combining this insertion with that of 6 nt at nt47 was also examined. As shown in Fig. 4, the 6-nt insertion at nt67, like that at nt47, had little or no effect on E307A amplification. However, the addition of 6 nt at both nt47 and nt67 reduced its amplification to below detectable levels, similar to the 12-nt NotI insertion at nt77. These results are consistent with the notion that a cis-acting sequence important for genome amplification lies downstream of this region [and presumably ends before nt20, the N protein start codon (Park et al., 1991)]. Its displacement downstream by 6 nt would appear to

be tolerated, but its displacement by 12 or 18 nt is not. The ability of E307 to tolerate a 42-nt insertion at nt72 remains unexplained, but could be due to the partial duplication of nt73-96, which contains the common sequence 5’ 73ctgAR GNYRR GYTCYN GANCYN96 (spac ing denotes hexamer phase). For comparison, the conserved sequence found at 77–96 nt from the 5’ ends of genomes and antigenomes (the BB box) can be similarly written as 77AR GRRCA GTYCN GACYYY96. Moreover, the possibility that a 42-nt insertion (but not a 12- or 18-nt insertion) can be tolerated at this position, independent of the sequence, cannot be excluded.

The le+ replication promoter and the rule of six

We also modified E307A by the insertion of 2 nt at nt47, and the compensatory deletion of 2 nt at nt67 (+2/−2) or the insertion of 4 nt at nt67 (+2/+4) (Fig. 5), so that the DI genome remained of hexamer length. E307A−2/0 was found to be amplified well, whereas E307A−2/4 was amplified very poorly. We at first assumed that E307A−2/4 contained an inadvertant nonhexamer deletion or insertion elsewhere which was responsible for its poor replication activity. However, the replacement of nt47-67 of
E307A with those of E307A*2/-4 indicated that the inactivating determinant was contained within these sequences, but we were unable to detect such a mutation in this region by resequencing the DNA (data not shown). Since the insertions of 6 nt at either nt47 or nt67 were well tolerated, and much of the sequence in between (nt52–55 and nt66–69) could be replaced without much effect (Fig. 4), these results were difficult to explain.

We therefore prepared the remainder of the E307A*2/-4 series, i.e., E307A*1/-5 to E307A*5/-1, as well as E307A*3/-3, and examined the effect of adding all possible permutations of 6 nt to this DI genome at nt47 or nt67 on its ability to be amplified. As shown in Fig. 5, the precise manner in which these 6 nt are added does have a strong effect on amplification activity. Whereas the addition of 6 nt entirely at nt47 (+6/0) or nt67 (0/+6) does not decrease amplificability relative to uninserted E307A, and the addition of 1 nt at one site and 5 nt at the other site had only a modest inhibitory effect, additions of +2/+4, +3/+3, or +4/+2 reduced amplification activity about 20-fold. In contrast, the addition of 2 or 3 nt at nt47 and the compensatory deletion of 2 or 3 nt at nt67 were relatively well tolerated (Fig. 5). As expected, the addition of 2 or 3 nt at nt47 without the compensatory deletion of 2 or 3 nt at nt67 inactivated these E307A derivatives and served as baseline controls (not shown). The addition of 6 nt, or altering the hexamer phase of nt47--67, individually, is thus tolerated, but the combination of both changes is not.

**DISCUSSION**

The cis-acting sequences which control paramyxovirus antigeneme synthesis are found at the 3' end of the (-) genome template, and their complement at the 5' end of the nascent (+) antigenome chain, where nucleocapsid assembly begins. We think it unlikely that the insertions on either side of the le'/N junction of DI-E307A exert their major effect on the assembly step of replication. To begin with, the first 42 nt in all our constructs are identical, and therefore the interaction of the first seven N subunits with the nascent antigenome RNA should be unaltered. We assume that this level of interaction is sufficient to drive nascent chain assembly, but this remains to be demonstrated. Further, it is difficult to see how the insertion of 6 nt at either nt47 or nt67 would have little effect on assembly, whereas the insertion of 12 or 18 nt at either site would have a strong effect, independent of their ability to form predicted RNA secondary structure (although the actual secondary structure in solution has not been examined). We think it more likely that the insertions exert their effect by affecting the ability of the polymerase to initiate RNA synthesis at the 3' end of the E307A (-) nucleocapsid. If so, our results suggest that the SeV polymerase interacts with (at least) two separate sequences for initiation of RNA synthesis.

One lies upstream of nt47 and would presumably include the conserved duodecamer (3' UGGUYU GUUCUC) at the ends and possibly extends to nt67 (Tapparel and Roux, 1996). The other element appears to be downstream of nt72 and may be the BB box [3' TTY CYGTY CARGNT CTGRRR as (-) RNA] (Blumberg et al., 1991). The le' / N junction which lies in between, on the other hand, does not appear to be critical by itself in genome amplification. Substitution of the terminal UUUU of the le' with either CCCG or GGGG, or of the 5' AGGGGU at the start of the N mRNA with 5' UCCCA (i.e., almost all of the sequence between the BglII and the Nsil sites), had little or no effect (data not shown). The 5' AGGGGU/UCCCA substitution (all transversions), moreover, eliminated mRNA synthesis from the E307A genome in vitro, whereas both of the 5' UUUU substitutions had little effect (data not shown). Amplification efficiency is thus unaffected by whether or not N/L mRNA synthesis is occurring from DI-E307 in our transfected cell system, consistent with previous results (Calain and Roux, 1995).

We have found that some mutations within nt1-77 are seemingly without effect by themselves, but are very poorly tolerated in combination. Two such examples have been found: the addition of 6 nt at nt47 or nt67, and the addition of 2 or 3 nt at nt67 and deletion of the same number of nucleotides at nt47 (thereby presumably changing the phase of the sequence between these sites relative to the N subunits of the template). E307A derivatives carrying each of these mutations individually (+6 at nt47, +6 at nt67, or the +2/-2 and +3/-3 constructs) are amplified almost as efficiently as E307A. The combination of two such mutations, however, are amplified very poorly, if at all, e.g., the +6/+6, and the +2/+4, +3/+3, and +4/+2 derivatives. These last constructs both add a total of 6 nt at nt47 and nt67 and change the phase of the intervening sequence. This situation, called synthetic lethality, is a prominent feature in the reverse genetic analysis of yeast. Mutants without phenotypes can still be linked with each other when the combination of the two is lethal (the obverse of a complementation group), presumably because their gene products have overlapping functions or function in the same pathway. In this view, the mutations listed above are partial loss of function mutations affecting the same pathway/function, the le' replication promoter. Apparently, a sequence at or downstream of nt2 can be displaced downstream by 6 nt (one N subunit) but not by 12 nt, or can be displaced downstream by 6 nt only if the hexamer phase of nt67--67 has not been altered, without strongly affecting amplification activity (Fig. 5). The combination of two such mutations, however, is synthetically lethal.

That the hexamer phase of nt47--67 may be important in determining genome amplification activity is of strong interest, because it is evidence that the rule of six operates at the level of the initiation of genome replication. This rule suggests that during the initiation
of genome synthesis, the viral polymerase interacts with its RNA template in the context of the N subunits, as each subunit is known to contact precisely 6 nt. The hexamer phase presumably results because the position of each nucleotide within a given subunit is different from each other within that subunit, but is the same between the same positions of different subunits. The cis-acting promoter sequence(s) is then viewed by the viral polymerase not only as a linear sequence, but as one in which the nucleotides are grouped as hexamers. Apparently, either part or all of nt$^{47-67}$ is recognized in the context of the N subunits, and this sequence can be displaced by 1, but not by 2, 3, or 4 nt positions downstream if a total of 6 nt have also been added in this region. Our results suggest that the polymerase is making multiple base contacts over at least 77 nt (13 N subunits, or one turn of the nucleocapsid helix, see below), some of which can be individually weakened without strongly affecting the initiation of antigenome synthesis. The fact that 6 nt can be inserted at at least two sites with little or no adverse effect (nt$^{47}$ and nt$^{67}$) suggests that the contact sites/promoter elements may be discontinuous (Fig. 6).

Our studies also suggest that only a limited amount of flexibility in the polymerase/template interactions can be accommodated for initiation. The displacement of the hexamer phase of nt$^{47-67}$ by 1 (but not by 2) nt positions is presumably well tolerated because the polymerase structure is sufficiently flexible for its contact site to move over by 1 nt position (but not by 2) relative to the template, to accommodate this displacement. Similarly, the insertion of 6 nt (but not 12 nt) is presumably well tolerated because, in this case, the template is sufficiently flexible for its contact site to move over by 1 N subunit (but not by 2) relative to the polymerase, for initiation. The SeV template is a helical assembly of N subunits and RNA, and this nucleocapsid coil is found in several discrete pitch states (Egelman et al., 1989; Charest et al., in preparation). These include relatively tight coils with pitches of 5.3 and 6.8 nm visible in negatively stained preparations (or 5.8 and 7.1 nm, in vitreous ice), as well as almost totally extended coils found in shadowed preparations (with a pitch of 37.5 nm) or in vitreous ice (with a pitch of 41 nm). The EM studies suggest that the SeV nucleocapsid is a dynamic structure, and the transitions between the different pitch states have been proposed to play a role in the packing of the helical nucleocapsid inside the spherical virion envelope, as well as in RNA synthesis (Heggeness et al., 1981; Egelman et al., 1989). The most extended form, in particular, has been proposed to be that which is copied by the viral P-L polymerase, and the interaction of the polymerase with the N:RNA template may induce the transition to the open form, at least locally. The initiation of RNA chains by bacterial and eucaryotic RNA polymerases on DNA templates occurs in three steps: (i) the reversible binding of RNA polymerase and ancillary proteins to promoter DNA to form the closed complex; (ii) the unwinding (isomerization) of a short stretch of DNA around the transcription initiation site to form the open complex, and the synthesis of aborted oligoribonucleotides; and (iii) promoter clearance and entry into the processive elongation phase (Eick et al., 1994). The nucleocapsid structural transitions mentioned above could also play a role in the initiation of genome chains and may explain why the promoter elements appear to be spread over such an extended region. The le$^+$ promoter also contains an element which renders it sensitive to C protein inhibition, adding further complexity to this control region (Cadd et al., 1996).

Finally, in a recent study using reporter CAT activity as a measure of the sum of SeV DI-analogue replication and transcription, Harty and Palese (1995) noted apparent exceptions to the rule of six. The deletion of 1 nt and its replacement elsewhere in the le$^-$ region was found
to be tolerated much better than the same operation in the \( \text{le}^+ \) region. These exceptions may very well only be apparent ones, since (i) the changes in the \( \text{le}^+ \) region may have affected the transcription promoter independently of the replication promoter, and (ii) the exact number of nucleotides in a chain cannot be the only important determinant of amplification efficiency, especially when changes (including hexamer phase changes of large blocks of sequence) are carried out in the leader regions, as these may destroy key promoter elements without altering total genome length. We have found the rule of six to be very stringent for SeV DI genome amplification in our transfected cell system. The only apparent exception to this rule that we have so far encountered are DI genomes containing the P gene mRNA editing site (Hausmann et al., 1996). When such DI genomes not of hexamer length are expressed in this system, their length is readjusted by the expansion (or contraction) of the purine run of the editing site during antigenome synthesis, and those of precise hexamer length are preferentially amplified. This apparent exception to the rule turns out, in fact, to be a demonstration of its stringency.

ACKNOWLEDGMENTS

We thank Joe Curran and Laurent Roux for numerous discussions. This work was supported by grants from the Swiss National Science Fund and the Human Frontier Science Program Organization.

REFERENCES


tion of mutant M proteins which interfere with virus particle budding. Virology 221, 159–171.