JOURNAL OF VIROLOGY, July 1997, p. 4944–4953 Vol. 71, No. 7
0022-538X/97/$04.00+0


# Analysis of the Gene Start and Gene End Signals of Human Respiratory Syncytial Virus: Quasi-Templated Initiation at Position 1 of the Encoded mRNA

LILI KUO, RACHEL FEARNS, AND PETER L. COLLINS*

*Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892-0720*

Received 15 October 1996/Accepted 18 March 1997

**The gene start (GS) and gene end (GE) transcription signals of human respiratory syncytial virus (RSV) strain A2 were analyzed in helper-dependent monocistronic and dicistronic minireplicons which were complemented by a standard RSV strain. The GS signal, which is the start site for mRNA synthesis, is highly conserved for the first nine genes: 3′-CCCCGUUUA(U/C) (negative sense). This conserved version of the signal was analyzed by "saturation" mutagenesis, in which all 10 positions, as well as one downstream and one upstream position, were changed one at a time into each of the other three nucleotides. Most of the positions appear to contribute to the signal: positions 1, 3, 6, 7, and, in particular, 9 were the most sensitive, whereas position 5 was relatively insensitive. The effect of nucleotide substitution in the first position of the signal was examined further by cDNA cloning and sequence analysis of the residual mRNA which was produced. For the two mutants examined (1C to U, and 1C to A), the site of initiation was unchanged. However, the mRNAs were dimorphic with regard to the assignment of the 5′-terminal nucleotide: two-thirds contained the predicted mutant substitution, and one-third contained the parental assignment. Intracellular minigenome contained only the mutant assignment, indicating that the heterogeneity was at the level of transcription by the RSV polymerase. This suggests that the templated mutant assignment at position 1 can sometimes be overridden by an innate preference for the parental assignment, a phenomenon which we dubbed quasi-templated initiation. The GS signal of the L gene, encoding the 10th RSV mRNA, contains three differences (3′-CCCUGUUUUA) compared to the conserved version. It was shown to be equal in efficiency to the conserved version. This was unexpected, since the saturation mutagenesis described above indicated that U in place of A at position 9 should be highly inhibitory. Instead, the A at position 10 of the L GS signal was found to be critical for activity, indicating that an essential A residue indeed was present in both versions of the GS signal but that its spacing differed. The GE signal, which directs termination and polyadenylation, has more sequence diversity in nature than does the GS signal. The naturally occurring GE signals of strain A2 were compared by their individual incorporation into a dicistronic minigenome. They were similar in the ability to produce translatable mRNA except in the cases of NS1 and NS2, which were approximately 60% as efficient.**

Human respiratory syncytial virus (RSV) is an enveloped RNA virus that is the leading viral agent of pediatric respiratory tract disease worldwide. RSV is classified in the *Pneumovirinae* subfamily of family *Paramyxoviridae* (9). Its genome is a single negative-sense strand of RNA of 15,222 nucleotides that encodes 10 subgenomic mRNAs (4, 9). The 10 major RSV proteins include the following: four nucleocapsid proteins, namely, the RNA-binding nucleocapsid N protein, phosphoprotein P, the large polymerase subunit L, and the transcription elongation factor encoded by the upstream translational open reading frame (ORF) of the M2 mRNA; three transmembrane envelope glycoproteins, namely, the fusion F protein, the attachment G protein, and the small hydrophobic SH protein; the inner virion matrix M protein; and two nonstructural proteins, NS1 and NS2. There is genetic evidence that an 11th protein is encoded by a second ORF in the M2 mRNA (7).

Transcription by RSV differs from that of rhabdoviruses and other paramyxoviruses in that it requires the transcription elongation factor M2 ORF1 protein in addition to N, P, and L (7, 15). Otherwise, RSV transcription and RNA replication seem to follow the general model for nonsegmented negative-strand RNA viruses based on vesicular stomatitis virus and Sendai virus (13, 16, 20, 21, 23). The polymerase enters the genome, which is encapsidated with N protein, in the extragenic leader region at the 3′ end. Transcription proceeds along the linear array of genes by a sequential, termination-reinitiation (stop-start) mechanism during which the polymerase remains template bound. This yields a series of subgenomic monocistronic mRNAs and, at least with the prototype viruses, a free leader RNA encoded by the leader region. RNA replication occurs when the polymerase somehow switches to a readthrough mode and synthesizes a complete positive-sense replicative intermediate, the antigenome, which also is encapsidated with N protein.

Stop-start transcription is directed by short conserved gene start (GS) and gene end (GE) signals which are found at the upstream and downstream boundaries, respectively, of each gene. The RSV GS signal is highly conserved; the first nine nucleotides are conserved exactly for 9 of the 10 genes of strain A2 (3′-CCCCGUUUAU/C), whereas that of the 10th gene, L, has three nucleotide differences: 3′-CCCUGUUUUA (see Fig. 1 and references 4, 5, 9, and 11). The first nucleotide of the GS signal is the mRNA start site. The GE motif of strain A2 is less highly conserved but conforms to the consensus 3′-UCA (A/U)U(-/U/G)(A/U)(A/U)(U/A)UUUU (positions having vari-

* Corresponding author. Mailing address: 7 Center Dr. MSC 0720, Bethesda, MD 20892-0720. Phone: (301) 496-3481. Fax: (301) 496-8312. E-mail: pcollins@atlas.niaid.nih.gov.

able assignments are in parentheses, with the assignments given in decreasing order of abundance, and the dash represents sequences which lack this position) (see Fig. 7 below and references 4, 5, and 9). The GE signal directs transcription termination and polyadenylation. Polyadenylation is thought to occur by reiterative copying of the short tract of U residues at the downstream end of the signal.

The poor growth of RSV in tissue culture and the instability of the virion have been obstacles to characterization of RSV RNA synthesis. Recently, however, systems have been developed for genetic analysis of RSV (2, 6, 7, 10, 15, 19, 20, 22, 24). One method involves short, cDNA-encoded, helper-dependent analogs of RSV genomic or antigenomic RNA (minigenome or miniantigenome, respectively). The minigenome or miniantigenome can be synthesized in vitro and complemented by transfection into RSV-infected cells, which is the approach used in this work. This has the advantage of taking place in the context of a permissive RSV infection free of any heterologous agent other than the in vitro-synthesized RNA analog and transfection reagent. Here, we used this method (i) to perform mutational analysis of the highly conserved version of the GS signal, (ii) to compare the conserved version of the GS signal with that of the L gene, and (iii) to compare the efficiencies of the naturally occurring GE signals.

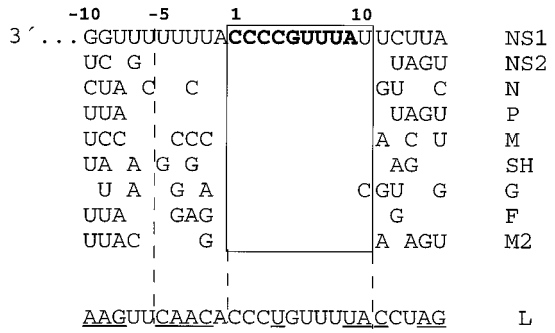


FIG. 1. Comparison of the naturally occurring GS signals and their flanking sequences of RSV strain A2, shown as negative sense. The top line shows the GS signal of the NS1 gene (the first nine nucleotides, which are conserved among the first nine RSV genes, are boldface, and nucleotides 1 to 10 are boxed) flanked by the 10 nucleotides which precede it (numbered −1 to −10) and the five nucleotides which follow it. The dashed lines are for ease of comparison. The corresponding sequences of all of the other genes except L are listed immediately below (in descending order according to their 3′ to 5′ position in the RSV gene order), with nucleotide assignments shown only when they differ from that of NS1. The L GS signal and flanking sequence are shown in full below, with nucleotide differences from NS1 underlined.

## MATERIALS AND METHODS

**Cells and virus.** RSV strain A2 of antigenic subgroup A was propagated in HEp-2 cells and stored at −70°C as tissue culture supernatants (typically $1 \times 10^7$ to $8 \times 10^7$ PFU per ml) adjusted to contain 50 mM HEPES (pH 7.5) and 100 mM magnesium sulfate. Transfections were performed in monolayer cultures of the adenovirus-transformed human embryonic kidney 293 cell line (10, 20).

**cDNAs.** cDNAs encoding the monocistronic RSV-CAT H15 minigenome or the dicistronic RSV-CAT-LUC minigenome were described previously (19, 20). The two negative-sense minigenomes, which are modeled after the A2 strain of antigenic subgroup A, are diagrammed below (see Fig. 2). Mutagenesis of RSV-CAT was performed by site-directed mutagenesis of single-strand phagemid DNA by the method of Kunkel et al. (18). The region of each mutation was confirmed by sequencing, and several independent clones of each mutant were analyzed in parallel for rescue to avoid possible artifacts due to spurious mutations elsewhere in the cDNA. Mutagenesis of RSV-CAT-LUC was done by PCR amplification using mutagenic primers spanning the *Nco*I and *Bam*HI restriction sites at the junction between the chloramphenicol acetyltransferase (CAT) and luciferase (LUC) genes (see Fig. 2) followed by restriction fragment replacement. All PCR-generated segments were sequenced completely.

**Minigenome synthesis and transfection.** Minigenome cDNA was linearized with *Hga*I or, in the case of constructs containing the ribozyme, with *Kpn*I, which cuts following the ribozyme and facilitates self-cleavage. Transcription by T7 RNA polymerase (Promega or Boehringer Mannheim) was done according to the procedures of the suppliers. RNA quality and quantity were confirmed by agarose gel electrophoresis and staining with ethidium bromide. 293 cells in six-well dishes ($1.5 \times 10^6$ cells per well) were infected by incubation for 45 min with 1 ml of RSV stock per well. The virus inoculum was replaced by 1 ml of Opti-MEM (Life Technologies) containing 12 μl of lipofectase (Life Technologies) and 1 to 5 μg of in vitro-synthesized minigenome RNA, an amount which was within the linear range of response (not shown). After 4 to 8 h of incubation, the cells were washed several times to remove residual input minigenome. Cells were harvested by scraping when cytopathic effect was extensive (24 to 30 h postinfection). Passage of clarified medium supernatants to fresh cells was performed as described previously (20).

**Enzyme assays.** Cell pellets were lysed by resuspension in 25 mM BICINE (*N*,*N*-bis-[2-hydroxyethyl]-glycine) (pH 7.5) containing 0.05% each Tween 80 and Tween 20, and the lysates were clarified by centrifugation. CAT activity was assayed by acetylation of [$^{14}$C]chloramphenicol as monitored by thin-layer chromatography (10, 20) and liquid scintillation counting of excised spots. Alternatively, CAT protein was measured by a commercial enzyme-linked immunoadsorption assay (CAT-ELISA; Boehringer Mannheim) (20). LUC activity was measured by using a commercial buffer system (Luciferase Assay System; Promega) and a Turner 20-TD luminometer (20).

**Northern (RNA) blot hybridization.** Total cellular RNA was purified with Trizol reagent (Life Technologies) (15). Northern blot analysis was performed by using a 1.5% agarose-formaldehyde gel and a nitrocellulose membrane (15) (in the case of Fig. 4) or with glyoxal denaturation, agarose gel electrophoresis, and a nylon membrane (20) (in the case of Fig. 6). In the case of Fig. 4B, the RNAs were first subjected to oligo(dT) chromatography by a batch method using oligo(dT) cellulose (15). Hybridization was performed with a $^{32}$P-labeled riboprobe synthesized in vitro by runoff transcription. CAT-specific negative- or positive-sense probe was synthesized from *Xba*I-digested RSV-CAT C2 or *Nco*I-digested C4 cDNA (15), respectively; negative-sense LUC probe was synthesized from pGEM-LUC plasmid (Promega); and negative- or positive-sense riboprobe specific to the RSV phosphoprotein (P) gene was synthesized from a P cDNA contained in pGem3 or pTM1, respectively (15).

**Rapid amplification of 5′ cDNA ends (5′-RACE), molecular cloning, and sequence analysis.** Intracellular RNA was isolated posttransfection and following the second passage, and mRNA was selected with biotin-labeled oligo(dT) and streptavidin-conjugated paramagnetic particles according to the directions of the supplier (PolyATtract System; Promega). The mRNA was subjected to reverse transcription (RT) under standard conditions using a negative-sense synthetic oligonucleotide, 5′-GCGGGAGGTAGATGAGATGTGACGAACGTGTACA TC, which annealed approximately 550 nucleotides from the 5′ end of the LUC mRNA. The cDNA was passed through a Glass Max (Life Technologies) spin column to remove unincorporated nucleotides, precipitated, and tailed with dG residues with terminal deoxynucleotide transferase. The reaction mixture was heated at 70°C, and the DNA was purified through a Glass Max spin column. The cDNA was amplified by PCR using $(dC)_{15}$ and a negative-sense oligonucleotide, 5′-TGTACATCGACTTGAAATCCCTGGTAATCCGTTTTAGAATC, that hybridized ~500 bp from the 5′ end of the LUC mRNA. The PCR conditions were 94°C (1 min), 51°C (1 min), and 72°C (2 min) for 25 cycles. This yielded the expected ~500-bp product, which was isolated by Gene Clean (BIO100) and cloned by using TA vector (Invitrogen). Two separate PCRs were done for each mutant, and the products were processed and cloned separately; approximately one-half of the sequenced cDNA clones were from each reaction mixture. As a control, mRNA was isolated from cells which received minigenome but not RSV. This did not yield a distinct PCR product and was not processed further. Intracellular minigenomes from the nonbinding fraction of the RNA from the second passage were subjected to RT using a positive-sense CAT-specific primer, 5′-A TGTCGGCAGAATGCTTA-3′, which hybridized approximately 90 nucleotides upstream of the intergenic region. The cDNA was passed through a Glass Max column, precipitated, and amplified by PCR using the same CAT-specific primer and the same negative-sense LUC-specific PCR primer as described above. The expected ~600-bp product was isolated and cloned as described above. Cells which had received minigenome but not RSV in the original transfection did not yield this PCR product. Dideoxynucleotide sequencing was performed with a Sequenase kit (U.S. Biochemicals) using a negative-sense LUC-specific primer which hybridized ~60 nucleotides from the 5′ end of the mRNA.

## RESULTS

**Mutational analysis of the GS signal.** The first nine nucleotides of the RSV GS signal are conserved exactly among the first nine genes of RSV strain A2, and eight of the first nine genes are identical at position 10: 3′-CCCCGUUUAU/C (Fig. 1). The GS signal of the L gene differs from this conserved sequence at three positions (Fig. 1) and is analyzed later. In addition, the GS signal of each of the 10 genes of strain A2 is preceded by an A residue, which suggested that it might have

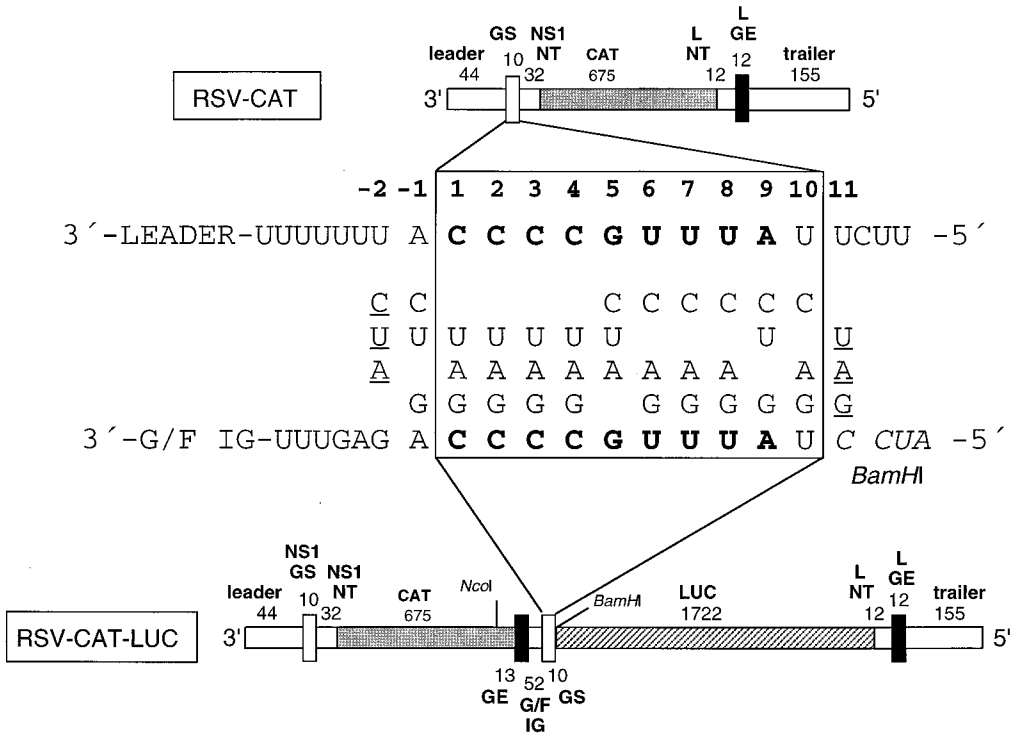


FIG. 2. Structures of the monocistronic minigenome RSV-CAT and the dicistronic minigenome RSV-CAT-LUC. The GS signal (nucleotides 1 to 10 are boxed, and conserved residues are boldface) and surrounding sequence are shown for each in negative sense. In between are the various substitutions which were analyzed in the GS signal of the CAT gene of RSV-CAT (which mimics that of the NS1 gene) and the GS signal of the LUC gene in RSV-CAT-LUC (which mimics that of the F gene). Mutations which were analyzed only in RSV-CAT-LUC (underlined) and the first four nucleotides of the *Bam*HI site which follows the LUC GS signal (italicized) are indicated. Abbreviations: NT, nontranslated; G/F IG, intergenic region between the G and F genes of strain A2.

been a nontranscribed part of the GS signal. The remainder of the flanking sequences are rich in U residues but otherwise lack obvious conserved sequences (Fig. 1). These sequence features also are very highly conserved in the human B subgroup and in bovine RSV (17, 25).

In a previous work (20), cDNAs encoding two negative-sense minigenome analogs of RSV genomic RNA were constructed (Fig. 2). RSV-CAT H15 is a monocistronic minigenome that contains a negative-sense copy of the CAT ORF under the control of a set of GS (conserved version) and GE transcription signals. RSV-CAT-LUC is a dicistronic minigenome modeled after RSV-CAT that contains in addition a negative-sense copy of the LUC ORF under the control of a separate set of GS (conserved version) and GE signals. The LUC gene is situated downstream of the CAT gene and separated from it by the natural 52-nucleotide G/F intergenic region.

The conserved version of the GS signal was subjected to "saturation" mutational analysis, in which each position was individually changed to each of the other three nucleotides. Analysis was performed first with the GS signal of the CAT gene of the monocistronic RSV-CAT minigenome. However, as described elsewhere (20), we found that approximately 5 to 10% of the transcription of the leader-proximal gene occurs by readthrough from the leader region independent of the GS signal, and this produced a high background. In contrast, transcription of the LUC gene in the second position in the dicistronic minigenome is completely dependent on its GS signal (20). Therefore, the complete set of nucleotide substitutions was reexamined, using the GS signal of the LUC gene of RSV-CAT-LUC. Analysis of the GS signal attached to two different genes in two different minigenomes provided independent confirmation.

The results of the saturation mutagenesis are shown in Table 1 and Fig. 3, which contain data of two complete representative experiments, one with RSV-CAT and one with RSV-CAT-LUC, in which each data point is an average of duplicate transfections. The duplicate samples usually differed by 10% or less. Also, the two minigenomes usually were in close agreement. One difference was that the RSV-CAT minigenome usually retained at least 5% CAT activity compared to that of the parent, which would be due to the above-mentioned readthrough from the leader region. In contrast, the most drastic mutations reduced LUC expression by the RSV-CAT-LUC minigenome to 0.2% or less compared to that of the parent. We considered a reduction of activity to 33% (indicated with a horizontal line in Fig. 3) or less to be significant.

Any substitution at positions 1, 3, 6, 7, or 9 reduced activity below 33% of that of the parent, and in particular, each substitution at position 9 was highly inhibitory. None of the substitutions at position 5 reduced activity below 33%. Positions 2, 4, and 8 were intermediate, with some of the substitutions being inhibitory. Thus, most of the positions of the GS signal influenced its activity. One of the natural GS signals of strain A2 has C in place of U at position 10, and this substitution had only a modest effect on activity. Two substitutions at position −1 and one at position 11 affected activity, but only the last one reduced activity to (slightly) below 33%. Thus, these positions influence GS activity but are not be considered part of the signal. This indicated that the conserved A residue at position −1 is not an important part of the signal.

TABLE 1. Saturation mutagenesis of the GS signal (conserved version) of the LUC gene in RSV-CAT-LUC or the CAT gene in RSV-CAT[a]

| Position[b] | % Activity following the indicated nucleotide substitution | | | |
|---|---|---|---|---|
| | C | U | A | G |
| −2 | 128.81 | 139.38 | 99.20 | |
| −1 | 44.80 (38.39) | 68.86 (60.31) | | 113.95 (100.0) |
| 1 | | 27.78 (21.20) | 10.80 (7.40) | 2.27 (16.30) |
| 2 | | 45.85 (45.80) | 18.59 (8.62) | 0.10 (8.90) |
| 3 | | 32.38 (16.50) | 19.44 (26.32) | 4.25 (10.50) |
| 4 | | 104.46 (107.7) | 15.60 (19.12) | 17.36 (12.13) |
| 5 | 37.15 (21.39) | 57.20 (60.19) | 104.04 (91.68) | |
| 6 | 4.30 (10.70) | | 7.19 (3.05) | 18.83 (21.89) |
| 7 | 11.25 (17.83) | | 0.57 (6.04) | 0.17 (6.23) |
| 8 | 61.24 (53.10) | | 8.66 (10.95) | 6.12 (11.05) |
| 9 | 0.27 (5.71) | 0.21 (4.35) | | 0.19 (5.77) |
| 10 | 54.57 (72.41) | | 23.66 (46.89) | 4.67 (15.83) |
| 11 | | 102.79 | 31.80 | 115.76 |

[a] Saturation mutagenesis (mutation of each position individually into each of the other nucleotides) was performed on the GS signal of the downstream LUC gene in RSV-CAT-LUC or (in parentheses) on the CAT gene in RSV-CAT. Activity is expressed as a percentage of that of the parental minigenome. Results of a single representative experiment are shown for each minigenome. Each data point is the mean of two samples; in general, the difference between duplicate samples was less than 10%. CAT activity was measured by the conversion of [$^{14}$C]chloramphenicol as measured by thin-layer chromatography.

[b] Sequence position relative to the conserved GS signal 3′-CCCCGUUUAU.

**Northern blot analysis of RNAs expressed from mutant RSV-CAT minigenomes.** It has previously been shown that ablation of the GS signal in RSV-CAT or RSV-CAT-LUC did not affect RNA replication (20). Under these conditions, expression of the marker enzymes as examined above would be an accurate measure of transcription. Nonetheless, we wanted to confirm that RNA replication was not affected in the present study and show that the levels of intracellular mRNA were consistent with the levels of enzyme expression. Therefore, intracellular RNA synthesized in response to RSV-CAT minigenomes bearing single nucleotide changes in the GS signal was isolated and examined by Northern blot hybridization with a riboprobe specific for positive-sense RNA. Figure 4 shows data for the parental RSV-CAT minigenome and for various mutants.

The RSV-CAT minigenome encodes two major species of positive-sense RNA: the miniantigenome and a polyadenylated subgenomic mRNA (20). In Northern blot analysis, the miniantigenome migrates as a discrete, somewhat larger species. Due to heterogeneity in the poly(A) tail, the mRNA appears as a diffuse band which in part migrates ahead of the miniantigenome and in part overlaps it. As shown in Fig. 4A, the amount of intracellular positive-sense RNA that accumulated in response to parental RSV-CAT and nine different mutants was approximately consistent with the levels of CAT expression shown in Table 1 and Fig. 3. The Northern blot analysis was somewhat more variable than the CAT assay, especially when the oligo(dT) selection step was included (see below), which probably reflects the greater number of steps in sample processing. Although the miniantigenome and mRNA were incompletely separated by gel electrophoresis, the RNA pattern of mutants which expressed low levels of CAT enzyme appeared to consist mainly of the somewhat larger, tightly banded miniantigenome, whereas mutants which expressed higher levels of CAT enzyme also had a diffuse mRNA component.

RNA representing the parental RSV-CAT genome and six different mutants was fractionated by oligo(dT) chromatography to separate the miniantigenome and mRNA and was analyzed by Northern blot hybridization (Fig. 4B). The accumulations of miniantigenome were similar for each of the minigenomes, confirming that RNA replication was not significantly affected by changes in the activity of the GS signal. The levels of poly(A)-selected mRNA approximately reflected the levels of CAT expression shown in Table 1 and Fig. 3. In this particular experiment, the recovery of mRNA appeared to be disproportionately low for mutant 1A (Fig. 4B, lane 2), which was not a consistent finding and represented experimental

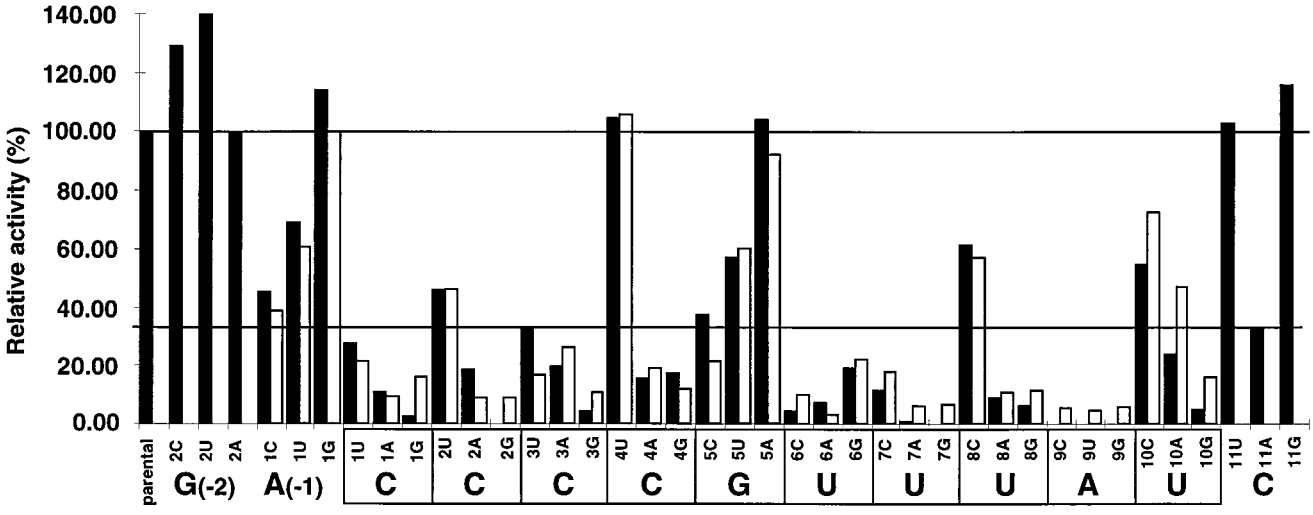


FIG. 3. Saturation mutagenesis of the conserved RSV GS signal, in which each position was changed individually into each of the other nucleotides. The parental sequence of the beginning of the LUC gene of RSV-CAT-LUC (at the bottom in large letters) (negative sense, 3′ to 5′) and the 10 nucleotides considered to constitute the GS signal (boxed) are shown. The various nucleotide substitutions at each position are shown along the *x* axis, and the percentage of activity compared to that of the parent is shown on the *y* axis (CAT expression by RSV-CAT [open boxes] and LUC expression by RSV-CAT-LUC [filled boxes] are indicated). This is a diagrammatic representation of the data in Table 1. Horizontal lines denote 100 and 33% activity, the latter being considered to be the threshold of a significant reduction.

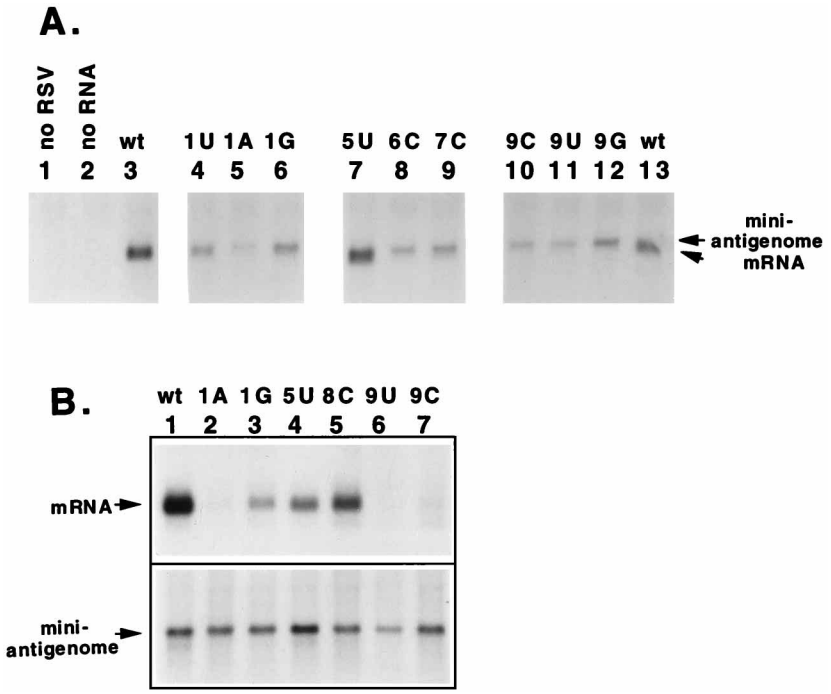


FIG. 4. Northern blot analysis of intracellular positive-sense RNA synthesized in RSV-infected cells in response to transfection with parental or mutant RSV-CAT minigenomes. (A) Uninfected (lane 1) or RSV-infected (lanes 2 to 13) cells were transfected with parental RSV-CAT RNA (lanes 1 and 3) or with the indicated RSV-CAT mutant RNA (lanes 4 to 13), identified by the GS signal position involved and its negative-sense assignment. Total intracellular RNA was isolated and subjected to Northern blot analysis using a negative-sense RSV-CAT riboprobe. (B) RNA was fractionated by oligo(dT) chromatography into bound (top panel) and unbound (lower panel) fractions which were analyzed by Northern blot hybridization using a negative-sense RSV-CAT riboprobe. wt, wild type.

variability. These data confirm that the mutations in the GS signal act at the level of transcription and leave RNA replication unaffected.

**Sequence analysis of the 5′ end of mRNAs encoded by minigenomes with a mutation at GS position 1.** Because the first four nucleotides of the conserved version of the GS signal are C residues (negative sense), it was possible that nucleotide substitution at position 1 would shift the start site to position 2. This was examined for two mutants of RSV-CAT-LUC, which had a U or an A in place of C at position 1 of the LUC GS signal. Intracellular mRNA produced by the parental minigenome or by mutant 1U or 1A was isolated posttransfection and also following passage 2. The mRNA was poly(A) selected, copied into cDNA by using an oligonucleotide primer specific to the LUC mRNA, and subjected to 5′-RACE. The predicted ~500-bp PCR product was observed as a distinct band by agarose gel electrophoresis (not shown). 5′-RACE of mRNA from control cells which had received minigenome but not RSV did not yield a detectable product (not shown). For each mRNA sample, duplicate PCRs were performed, and the cDNA from each was recovered and cloned separately. Approximately equal numbers of clones from each reaction were sequenced to determine the structure of the 5′ mRNA end.

Sixteen to 33 cDNA clones of mRNA were sequenced each for the parent and the two mutants (Fig. 5, top and middle panels; Table 2). A second, independent experiment of approximately equal size was performed and yielded results essentially the same as those shown in Fig. 5 and Table 2 (not shown). The sequencing ladders in Fig. 5 (which are negative sense) show that sequence of the 5′ end of the parental LUC mRNA is 5′-(G)GGGGC... (positive sense). The presence of the additional, fifth G residue indicated in parentheses was observed previously when the 5′ end of authentic RSV F mRNA (which is the model for the LUC GS signal [Fig. 2]) was sequenced by primer extension (8) and might be due to copying of the mRNA cap (1) during RT. This extra nucleotide was present in 86% of the clones, and its frequency was the same for the parent and the two mutants. The position in the sequencing ladder above this base, which should be a G contributed by the homopolymer tail added during cDNA synthesis, frequently had a band in the A lane, or in both the A and the G lanes, or across all lanes. This was variable and appeared to be an artifact of the sequencing reaction.

Examination of the 46 cDNAs representing the 1U and 1A mutants showed that the mRNA start site remained invariant at position 1 of the GS signal despite the changes in nucleotide assignment. Unexpectedly, the 1U and 1A mutants each exhibited dimorphism in the structure of cDNA representing the 5′ mRNA end. The predominant sequences (positive sense) deduced for the 1U and 1A mutants were, respectively, 5′-(G) AGGGC (20 of 30 cDNA) and 5′-(G)UGGGC (9 of 16 clones) (Fig. 5, top and middle panels, and Table 2; note that nucleotide assignments in Table 2 are negative sense). These contained the complement of the point mutation but otherwise were the same as parental LUC mRNA. Remarkably, the remainder of the clones in each case contained the wild-type assignment and were identical to parental LUC mRNA. Sequencing ladders of these wild-type sequences from mutant mRNA are not shown because they are essentially identical to the wild-type ladders in Fig. 5. One possibility was that there was sequence heterogeneity in the intracellular minigenome template. However, sequence analysis of minigenomes isolated from the second passage showed that each template contained only the predicted assignment (Fig. 5, bottom panel). We also considered the possibility that the plasmid preparations for the mutants might be contaminated by the parent. This seemed

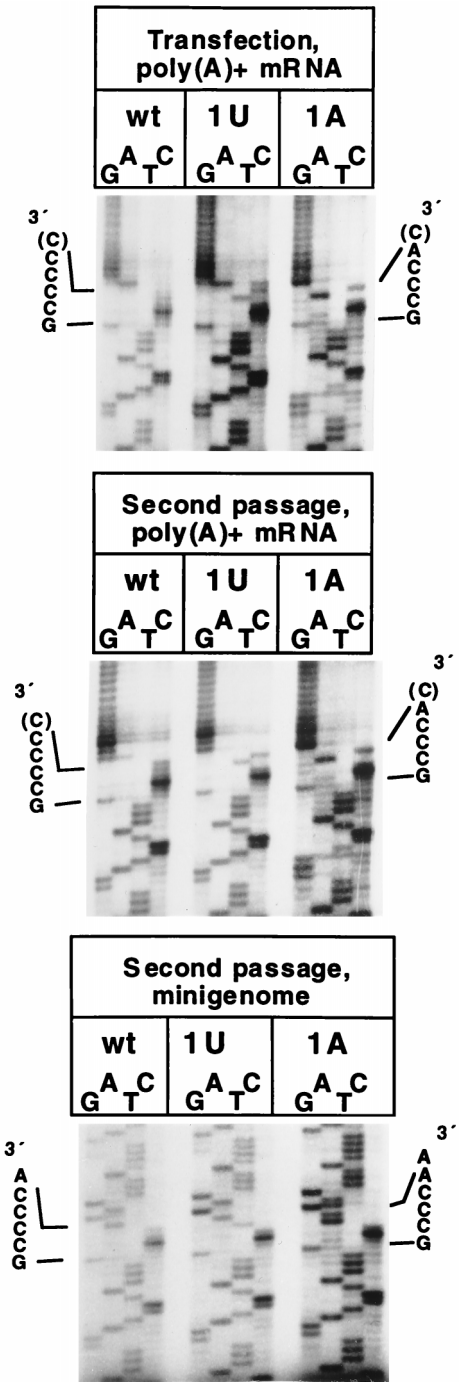


FIG. 5. Sequence analysis of the 5′ end of LUC mRNA synthesized from parental RSV-CAT-LUC minigenome (wild type [wt]) or by mutants containing a single nucleotide substitution at position 1 in the LUC GS signal (1U and 1A). RSV-infected cells were transfected with minigenome and subjected to two serial passages. Poly(A)-selected mRNA from the transfection (top panel) and from passage 2 (middle panel) was subjected to 5′-RACE, and intracellular minigenome from passage 2 was subjected to RT-PCR (bottom panel). The cDNAs were cloned, and individual clones were analyzed by nucleotide sequencing. The deduced negative-sense sequence of the mRNA or minigenome is shown on the left for the wt and on the right for the 1A mutant. In the top and middle panels, an additional nucleotide which might represent copying of the 5′ mRNA cap is indicated in parentheses.

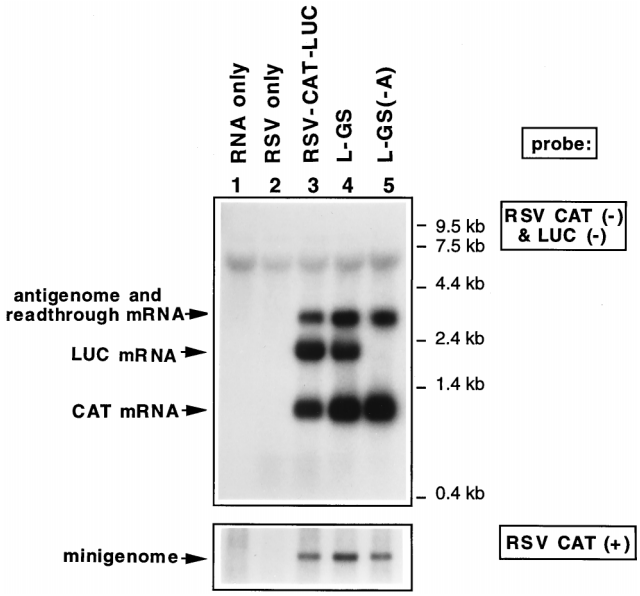


FIG. 6. Northern blot analysis of intracellular RNAs encoded by RSV-CAT-LUC parental minigenome (lane 3) or by mutants in which the GS signal of the downstream LUC gene has been replaced with that of the L gene (L-GS, lane 4) or that of the L gene except that the A residue at position 9 had been deleted [L-GS(-A), lane 5]. Replicate blots were analyzed by hybridization with a mixture of negative-sense RSV-CAT and LUC riboprobes (upper panel) or a positive-sense RSV-CAT riboprobe (lower panel). The positions and kilobase lengths of RNA molecular size markers are shown for the upper panel. The band that migrated between the 4.4- and the 7.5-kb markers and is present in all lanes is 28S rRNA.

unlikely; whereas mutagenesis methods employing heteroduplex intermediates can give rise to mixed plasmid populations, the method used here with the RSV-CAT-LUC constructs involved PCR and restriction fragment swapping followed by cloning. Nonetheless, we examined this directly by transforming the plasmid DNA preparations into *Escherichia coli* and

TABLE 2. Number of cDNA clones containing the indicated assignment (negative sense) as the first nucleotide of the GS signal, determined from intracellular mRNA or minigenome

| RNA analyzed | No. of clones with the indicated nucleotide at position 1/no. tested | | |
|---|---|---|---|
| | C (parental) | U | A |
| mRNA of transfection[a] | | | |
| Parental 1C | 17/17 | | |
| Mutant 1U | 6/16 | 10/16 | |
| Mutant 1A | 2/6 | | 4/6 |
| mRNA of 2nd passage[a] | | | |
| Parental 1C | 16/16 | | |
| Mutant 1U | 4/14 | 10/14 | |
| Mutant 1A | 5/10 | | 5/10 |
| Minigenome of 2nd passage[b] | | | |
| Parental 1C | 6/6 | | |
| Mutant 1U | | 6/6 | |
| Mutant 1A | | | 6/6 |

[a] Intracellular poly(A)-selected mRNA was subjected to RT with a negative-sense CAT-specific oligonucleotide, the cDNA was tailed with oligo(dC), and PCR was performed followed by cDNA cloning and nucleotide sequencing.

[b] Intracellular minigenome was reverse transcribed with a positive oligonucleotide specific to the CAT gene, amplified by PCR, cloned, and sequenced.

TABLE 3. Comparison of the conserved versus L-specific versions of the GS signal and importance of the A residue at position 10 in the L GS signal

| LUC GS signal in RSV-CAT-LUC[a] | CAT activity | | LUC activity[d] | | LUC/CAT ratio[e] | | Normalized ratio (%)[f] | | |
|---|---|---|---|---|---|---|---|---|---|
| | Expt I[b] | Expt II[c] | Expt I | Expt II | Expt I | Expt II | Expt I | Expt II | Mean (I + II) |
| GS consensus | 274.0 | 9.36 | 110.3 | 213.1 | 0.403 | 22.69 | 100.0 | 100.0 | 100.0 |
| L GS | 265.7 | 10.22 | 125.4 | 186.9 | 0.471 | 18.28 | 116.87 | 80.54 | 98.71 |
| L-GS(-A) | 267.3 | 10.34 | 10.69 | 3.175 | 0.04 | 0.31 | 9.92 | 1.35 | 5.64 |
| Control: RNA only[h] | <0.1 | <0.1 | 0.001 | 0.000 | ND[g] | ND | ND | ND | ND |
| Control: RSV only | <0.1 | <0.1 | 0.002 | 0.004 | ND | ND | ND | ND | ND |

[a] Structures of GS for the LUC gene in RSV-CAT-LUC:

| | −1 1 | 9 | 10 |
|---|---|---|---|
| GS consensus | 3′ A | CCCCGUUUA | U/C |
| L GS | A | CCCUGUUUU | A |
| L-GS(−A) | A | CCCUCUUUU | C |

[b] In experiment ICAT was measured by enzyme-linked immunosorbent assay (ELISA) and is expressed as picograms per 6,000 cells.
[c] In experiment II, CAT was measured by acetylation of [$^{14}$C]chloramphenicol and is expressed as percent conversion in a reaction using lysate from $7.5 \times 10^4$cells.
[d] Luciferase light units per 6,000 cells.
[e] Ratio of LUC and CAT activities, as arbitrary units.
[f] Each experiment was normalized separately with the GS consensus taken as 100%.
[g] ND, not determined.
[h] RSV-CAT-LUC containing GS consensus as the LUC GS signal.

sequencing the GS sequence of 36 randomly selected clones each for the 1U and 1A plasmid preparations. All 72 clones contained the appropriate mutant assignment, indicating an absence of contamination by the parent plasmid.

**GS signal of the L gene.** As shown in Fig. 1, the GS signal of the L gene of strain A2 differs from that of the other nine genes at three positions, 4, 9, and 10. These differences are conserved between the two antigenic subgroups (17). As described above (Table 1 and Fig. 3), mutation of the conserved version of the GS signal to contain the L-specific assignment at position 4 (U instead of C) had no effect on transcription, and the difference at position 10 (A instead of U or C) was associated with a modest reduction in transcription. In contrast, mutation of position 9 to contain the L-specific assignment (U instead of A) rendered the conserved version of the signal inactive (Table 1 and Fig. 3). It therefore was of interest to compare the activities of the conserved GS signal and the L-specific version. The conserved GS signal of the LUC gene of RSV-CAT-LUC was replaced with the L-specific version. This mutant was equivalent to the parent in the efficiency of synthesis of LUC mRNA (Table 3 and Fig. 6), indicating that the conserved and L-specific versions of the GS signal were equivalent in activity.

This suggested that some feature of the L GS signal compensates for the absence of A at position 9. One possibility was that the A residue at position 10 of the L GS signal might take its place. If so, then position 10 would be highly sensitive to deletion or substitution. To test this, the A residue at position 10 was deleted to generate the mutant L-GS(-A). This deletion had the effect of moving a C residue into position 10, which corresponds to a naturally occurring assignment in the conserved version of the GS signal, and position 11 remained C. Removal of 10A reduced LUC expression by an average of nearly 20-fold (Table 3) and ablated the accumulation of LUC mRNA (Fig. 6, lane 5). Thus, the A residue at position 10 appeared to be critical for the activity of the LUC GS signal, comparable to the A residue at position 9 of the conserved version of the GS signal.

**Comparison of naturally occurring GE signals of RSV strain A2.** Whereas the GS signal of strain A2 is highly conserved among the different genes, the GE signal shows greater variability (Fig. 7). To determine whether this has significance for gene transcription, each of the 10 naturally occurring GE signals of strain A2 was inserted individually into the LUC gene of the dicistronic minigenome RSV-CAT-LUC (Fig. 7). It was previously shown that replacement of the GE signal with a random sequence ablated polyadenylation and termination, leading to transcriptional readthrough (20). In the case of the LUC gene of RSV-CAT-LUC, this drastically reduced the accumulation of LUC mRNA, presumably because the nonpolyadenylated mRNA is unstable, and reduced the expression of LUC enzyme more than 200-fold to background levels (20). There was no effect on RNA replication. Therefore, measurement of LUC activity in the RSV-CAT-LUC minigenome is a sensitive measure of the activity of the GE signal.

CAT and LUC enzyme activities were measured for each of the RSV-CAT-LUC minigenomes representing the naturally occurring GE signals of strain A2. Four separate experiments were performed in which each minigenome was represented in multiple samples (Table 4). Transcription of the upstream CAT gene was shown previously to be independent of transcription of the downstream LUC gene (20) and therefore served as an internal standard. Therefore, within each experiment, the average of the LUC activity for each minigenome was divided by the average of its CAT activity. This provided a value in arbitrary units which could be used to directly compare the expression of LUC by the various minigenomes within each experiment. The value for each experiment was then normalized relative to that of RSV-CAT-LUC(L-GE) as 1.00. Finally, the normalized values for the four experiments were averaged (Table 4) and depicted graphically (Fig. 8). This showed that the various GE signals were nearly identical in activity except for those of NS1 and NS2, which were approximately 60% as efficient.

## DISCUSSION

The GS and GE transcription signals of RSV strain A2 were examined by using a system based on transfection of in vitro-synthesized minigenome RNA into RSV-infected cells. The conserved version of the GS signal, for which the first nine nucleotides are identical among the first nine genes of strain A2, was analyzed by saturation mutagenesis. This involved

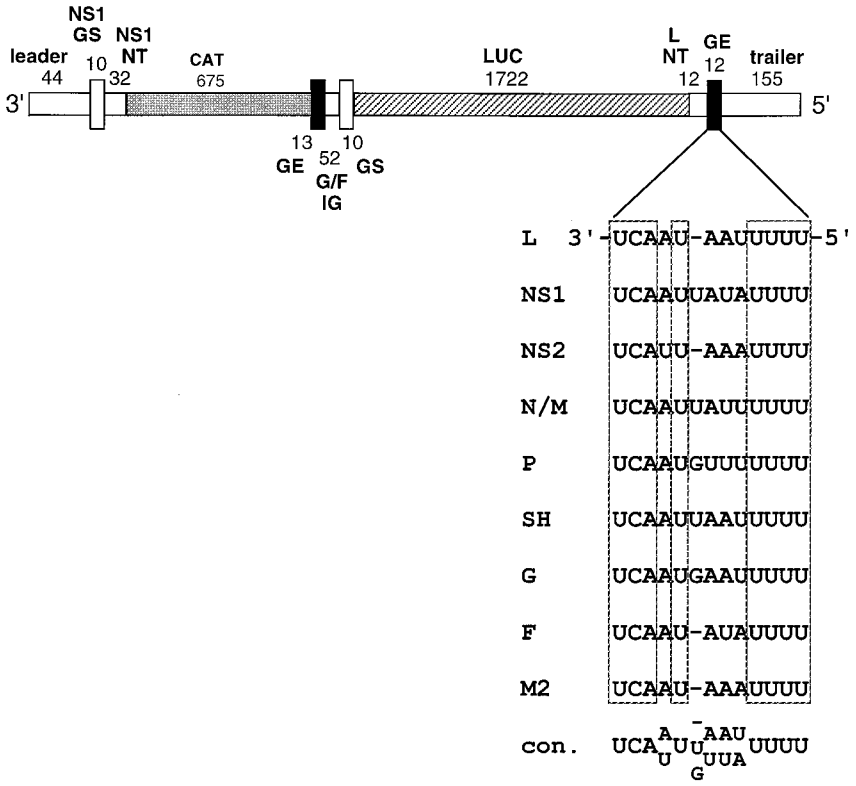


FIG. 7. Insertion of the various naturally occurring GE signals of strain A2 into the LUC gene of the RSV-CAT-LUC dicistronic minigenome. In the consensus sequence at the bottom, assignments at variable positions are given in order of decreasing frequency. NT, nontranslated; con., consensus.

changing each sequence position in turn into each nucleotide assignment. Differences between the conserved version and the L-specific version of the GS signal also were examined. The GE signal is more divergent among the different genes than is the GS signal, and the activities of the various GE signals of strain A2 were compared.

It has previously been shown that the GS signal is required for transcription of its gene (20). Mutational analysis of the conserved version of the GS signal was performed in parallel for the GS signal of the CAT gene of RSV-CAT and the GS signal of the LUC gene of RSV-CAT-LUC. The results obtained with the two minigenomes generally were in accord and showed that most of the positions of nucleotides 1 to 10 of the signal were important for its activity. Position 5 alone was tolerant of most substitutions, and several other positions tolerated certain substitutions. Substitution by G always reduced activity well below 33%, whereas the effect of substitution by U, A, or C was variable. Any substitution at position 9 ablated

TABLE 4. Comparison of the 10 naturally occurring GE signals of strain A2 by their individual insertion into the LUC gene of the dicistronic RSV-CAT-LUC minigenome[a]

| GE[b] | LUC/CAT ratio (mean ± 1 SD)[c] in expt: | | | | Normalized ratio[d] in expt: | | | | Mean ± 1 SD (I + II + III + IV) |
|---|---|---|---|---|---|---|---|---|---|
| | I ($n$ = 3)[e] | II ($n$ = 3)[e] | III ($n$ = 3)[f] | IV ($n$ = 3)[f] | I | II | III | IV | |
| NS1 | 0.32 ± 0.017 | 0.27 ± 0.060 | 13.53 ± 1.55 | 12.65 ± 1.23 | 0.50 | 0.55 | 0.64 | 0.62 | 0.58 ± 0.07 |
| NS2 | 0.34 ± 0.014 | 0.29 ± 0.024 | 15.41 ± 2.63 | 13.09 ± 1.60 | 0.54 | 0.60 | 0.73 | 0.64 | 0.60 ± 0.08 |
| N (M) | 0.60 ± 0.063 | 0.57 ± 0.099 | 19.47 ± 4.22 | 20.65 ± 0.98 | 0.94 | 1.18 | 0.92 | 1.00 | 1.01 ± 0.12 |
| P | 0.59 ± 0.047 | 0.46 ± 0.025 | 20.25 ± 4.58 | 21.37 ± 5.76 | 0.92 | 0.95 | 0.96 | 1.04 | 0.97 ± 0.05 |
| SH | 0.53 ± 0.024 | 0.42 ± 0.020 | 20.51 ± 4.09 | 18.74 ± 3.38 | 0.83 | 0.87 | 0.97 | 0.91 | 0.90 ± 0.06 |
| G | 0.57 ± 0.031 | 0.44 ± 0.025 | 23.81 ± 2.53 | 18.54 ± 0.59 | 0.89 | 0.92 | 1.13 | 0.90 | 0.96 ± 0.11 |
| F | 0.54 ± 0.085 | 0.43 ± 0.020 | 24.01 ± 2.05 | 18.95 ± 1.17 | 0.85 | 0.90 | 1.14 | 0.92 | 0.95 ± 0.13 |
| M2 | 0.58 ± 0.024 | 0.41 ± 0.011 | 24.79 ± 1.63 | 17.74 ± 0.37 | 0.91 | 0.85 | 1.18 | 0.86 | 0.95 ± 0.15 |
| L | 0.64 ± 0.036 | 0.48 ± 0.042 | 21.10 ± 4.06 | 20.57 ± 2.47 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 ± 0.00 |

[a] The CAT and LUC activities from control cells which received only RSV or RNA gave background readings and are not shown.
[b] GE signal placed on the downstream end of the LUC gene.
[c] For each sample, the LUC activity was divided by the CAT activity, the latter serving as an internal standard. The ratio is an arbitrary unit for comparison within an individual experiment; the magnitude of this arbitrary value varied greatly, depending on whether the experiment involved CAT measurement by ELISA or chloramphenicol acetylation.
[d] Each experiment was normalized relative to the L GE signal as 1.00.
[e] CAT was measured by CAT ELISA (Boehringer Mannheim Biochemicals), represented in picograms per 6,000 cells.
[f] Percent acetylation of [$^{14}$C]chloramphenical substrate by lysate from 7.5 × $10^4$ cells.

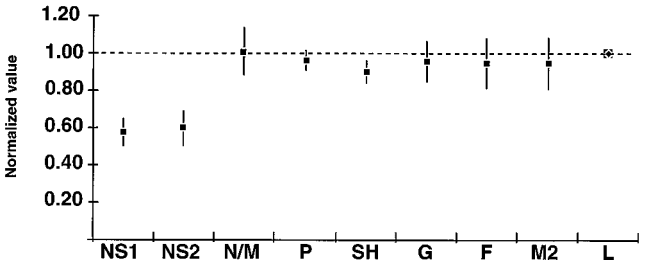


FIG. 8. Comparison of the activities of the various naturally occurring strain A2 GE signals when placed individually in the LUC gene of the RSV-CAT-LUC minigenome (see Table 4). The value for the L GE signal is unity because it was the standard for normalization. Means (boxes) ± 1 standard deviation (vertical lines) are shown.

activity; the ~5% residual activity observed for the RSV-CAT mutants at this position is due to the above-mentioned GS-independent readthrough transcription from the leader region (20). Mutations in the −1, +1, or +2 position had a moderate effect or no effect, indicating that the GS signal was contained within the 10-nucleotide conserved motif. This is consistent with previous results (20) which showed that authentic sequence could be resected from around the 10-nucleotide motif without loss of function.

GS motifs have been described for a number of the genes of strain 18537, representing human antigenic subgroup B (strain A2 represents subgroup A), for the closely related bovine RSV, and for the more distantly related pneumonia virus of mice (3, 17, 25). Of those GS motifs which have determined for strain 18537 and bovine RSV, only two differences were noted relative to the A2 strain. One difference involved the L GS signal of the bovine strain and is described below, and the second involved the SH gene of both the subgroup B and bovine strains, in which in each case position 5 of the GS signal is A instead of G. As shown in Fig. 3 and Table 1, placing this substitution in the conserved form of the GS signal had no effect on its activity. The GS signal of pneumonia virus of mice differs from that of RSV strain A2 at three positions: position 1 is U instead of C, position 4 is U instead of C (as in the human RSV L GS signal), and position 8 is C instead of U. In each case, this difference represents the most benign substitution possible (Fig. 3 and Table 1), and the substitutions at positions 4 and 8 did not reduce activity below 33%. These comparisons show that the nucleotide differences that exist between these different viruses tend to be those which are the least deleterious to the activity of the conserved version of the signal. This probably represents divergent evolution constrained by a bias against mutations which greatly reduce function.

Replacement of the native assignment of C at GS position 1 with U or A left 27.8 or 10.8% residual LUC expression (Table 1). This indicated that there was not an absolute requirement for the authentic assignment. This was contrary to our expectation that the first position might be particularly sensitive to substitution, and it was somewhat surprising that the 1A mutant retained detectable expression since polymerases usually initiate mRNA with a purine. Since positions 2, 3, and 4 are C residues, like position 1 of the parental signal, it was possible that the residual transcription by the 1U and 1A mutants might be due to initiation at one of these sites. Therefore, the 5′ end of the mRNA produced by the RSV-CAT-LUC minigenome bearing the 1U or 1A mutation was amplified by 5′-RACE, cloned, and sequenced.

Analysis of numerous cloned cDNAs indicated that the polymerase always initiated at position 1, even when the nucleotide assignment at that position was changed. Thus, the position of the start site appears to be a strictly controlled feature of the GS signal. Surprisingly, however, only two-thirds of the mRNA cDNAs contained the mutant assignment at the first position: the remaining third contained the parental assignment. This was not due to heterogeneity of the intracellular minigenome template, as confirmed by RT-PCR, cloning, and sequence analysis. Therefore, the mutant coding assignment specified by the template for the first nucleotide of the mRNA apparently can be overridden. We dubbed this phenomenon quasi-templated initiation, and to our knowledge it has not been described previously.

It might be that the RSV polymerase is preloaded with the parental initiating nucleotide and thus has an inherent bias for that assignment. Another possibility is that "correction" occurs in *trans* by polymerases which initiate on helper genomes and jump to the minigenome with nascent transcript in tow. This can be tested using a previously described transcription-replication system that is reconstituted entirely from transfected plasmids and is free of helper virus and free of mRNAs containing the pertinent 5′-end sequence (7, 15). A third possibility is that the fraction of mRNA containing the authentic assignment results from initiation at position 2, 3, or 4 followed by a slippage-realignment mechanism, as has been described for certain RNA viruses such as tacaribe virus (14). This mechanism might not be part of the normal mechanism of initiation but might be an inherent capability of the polymerase that comes into play with certain mutants. We also noted that most of the sequencing ladders of mRNA-derived cDNA contained an additional C residue adjacent to the 5′-terminal nucleotide. This was previously observed during sequence analysis of the 5′ end of the RSV F mRNA (8) and is thought to be a result of copying the cap (1) during RT. The presence of this additional nucleotide in each of the sequencing ladders suggests that the mutant versions of the mRNA were capped. That the CAT mRNA was capped in each situation also is suggested by its evident translatability: there is concordance between the levels of RNA and expressed CAT enzyme. It is noteworthy that the mRNAs under examination were synthesized by RSV-encoded polymerase in the absence of any heterologous factors other than the in vitro-synthesized RNA and the transfection reagent. Thus, the structures under examination should be authentic.

The L GS signal differs from the conserved version at positions 4, 9, and 10. On the basis of the mutational analysis described above, the difference at position 4 would be insignificant, and the difference at position 10 would have only a moderate effect. In contrast, substitution of A by U at position 9 of the conserved version of the GS signal rendered it inactive. Direct comparison of the conserved versus the L-specific versions of the GS signal showed that the two were equivalent in activity. This suggested that some other feature compensates for the difference at position 9. We speculated that this feature might be the A residue at position 10 in the L GS signal. If this residue is the counterpart of the 9A residue in the conserved version, then it too should be highly sensitive to removal or substitution. This indeed proved to be the case; its removal reduced transcription of the attached LUC gene to levels which were undetectable by Northern blotting. Removal of the 10A residue had the effect of moving a C residue to position 10 (while position 11 remained C), but that probably was not significant because a C residue occurs naturally at position 10 in the G GS signal, and its placement in position 10 by mutational analysis also was not inhibitory. Thus, the 9A and 10A residues in the conserved and L-specific versions of the GS signal, respectively, are each critical to the activity of the respective signal and might be functional counterparts. If so, this

illustrates that, while the polymerase always begins synthesis at position 1 of the signal, there is flexibility in the position of the critical A residue. Interestingly, the L GS signal of the bovine strain has an apparent deletion of position 1, such that the A at position 10 of the human strains is at position 9 of the bovine strain and corresponds directly with 9A of the conserved version of the GS signal (25).

The naturally occurring RSV GE signals exhibit some sequence diversity: those of strain A2 contain the conserved pentamer UCA(A/U)U followed by four variable positions, followed by four U residues. Evaluation of the 10 GS signals of strain A2 by their individual insertion into the LUC gene of RSV-CAT-LUC showed that eight of the signals, namely, those of the N, P, M, SH, G, F, M2, and L genes, were indistinguishable in activity. However, the GE signals of the two promoter-proximal genes, NS1 and NS2, were 60% as active as the others. Reduced efficiency of termination and polyadenylation for these two promoter-proximal genes would lead to increased production of their associated readthrough transcripts. Readthrough mRNAs are synthesized during natural RSV transcription at a frequency of 0.1 relative to the monocistronic mRNAs (12, 19). Specifically, this would increase the synthesis of the readthrough mRNAs NS1-NS2, NS1-NS2-N, and NS2-N at the expense of the monocistronic NS1, NS2, and N mRNAs. Indeed, readthrough mRNAs involving the first three RSV genes are very prominent in natural infection (12). This shift in favor of readthrough mRNAs for the first three genes would decrease the synthesis of the NS2 and N proteins, since they would not be efficiently translated from downstream ORFs in readthrough mRNAs. Thus, the structures of the GE signals might have a modest influence in determining the relative levels of synthesis of the NS2 and N proteins. The NS2 GE signal is the only one which contains U rather than A at the otherwise conserved position 5, which might account for its reduced activity. Position 5 is U also in six GE signals sequenced to date for strain 18537 (17). The NS1 GE signal does not contain any obvious features which might account for reduced activity, although it is the only signal in strain A2 or among the six sequenced to date for strain 18537 which both is 13 nucleotides long and contains only 4 U residues at the downstream end.

The next logical step for these studies will be to evaluate individual mutations by their introduction into infectious virus (6). Mutational analysis of transcription signals in a "wild-type" recombinant virus would be complicated by the likelihood it would alter the level of expression of individual genes and alter the growth properties of the virus. However, a recombinant RSV bearing the CAT ORF as an additional, foreign gene was recently described (2). Since the CAT gene is heterologous, its product would not play a role in the growth of this recombinant. Alteration of the efficiency of the CAT GS signal should have minimal effect on transcription of the other genes. We note that this probably would not be the case for the GE signal, since the activity of this signal would be important for initiation at the next downstream RSV gene (20). It also will be interesting to determine whether the introduction of certain of the GS mutations or GE substitutions described here into one or more RSV genes would have an attenuating effect which might be exploited for the development of live attenuated RSV as a vaccine.

## ACKNOWLEDGMENTS

We thank Myron Hill and Ena Camargo for technical assistance and Robert Chanock, Brian Murphy, Mark Peeples, Prabha Atreya, and Alexander Bukreyev for helpful comments on the manuscript.

## REFERENCES

1. **Barik, S.** 1993. The structure of the 5′ terminal cap of the respiratory syncytial virus mRNA. J. Gen. Virol. **74:**485–490.
2. **Bukreyev, A., E. Camargo, and P. L. Collins.** 1996. Recovery of infectious respiratory syncytial virus expressing an additional, foreign gene. J. Virol. **70:**6634–6641.
3. **Chambers, P., D. A. Matthews, C. R. Pringle, and A. J. Easton.** 1990. The nucleotide sequences of intergenic regions between nine genes of pneumonia virus of mice establish the physical order of these genes in the viral genome. Virus Res. **18:**263–270.
4. **Collins, P. L.** 1991. The molecular biology of human respiratory syncytial virus (RSV) of genus Pneumovirus, p. 103–162. *In* D. W. Kingsbury (ed.), The paramyxoviruses. Plenum Publishing Corp., New York, N.Y.
5. **Collins, P. L., L. E. Dickens, A. Buckler-White, R. A. Olmsted, M. K. Spriggs, E. Camargo, and K. V. W. Coelingh.** 1986. Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order. Proc. Natl. Acad. Sci. USA **83:**4594–4598.
6. **Collins, P. L., M. G. Hill, E. Camargo, H. Grosfeld, R. M. Chanock, and B. R. Murphy.** 1995. Production of infectious human respiratory syncytial virus from cloned cDNA confirms an essential role for the transcription elongation factor from the 5′ proximal open reading frame of the M2 mRNA and provides a capability for vaccine development. Proc. Natl. Acad. Sci. USA **92:**11563–11567.
7. **Collins, P. L., M. G. Hill, J. Cristina, and H. Grosfeld.** 1996. Transcription elongation factor for respiratory syncytial virus, a nonsegmented negative strand RNA virus. Proc. Natl. Acad. Sci. USA **93:**81–85.
8. **Collins, P. L., Y. T. Huang, and G. W. Wertz.** 1984. Nucleotide sequence of the gene encoding the fusion (F) glycoprotein of human respiratory syncytial virus. Proc. Natl. Acad. Sci. USA **81:**7683–7687.
9. **Collins, P. L., K. McIntosh, and R. M. Chanock.** 1995. Respiratory syncytial virus, p. 1313–1351. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), Virology, 3rd ed. Raven Press, New York, N.Y.
10. **Collins, P. L., M. A. Mink, and D. S. Stec.** 1991. Rescue of synthetic analogs of respiratory syncytial virus genomic RNA and effect of truncations and mutations on the expression of a foreign reporter gene. Proc. Natl. Acad. Sci. USA **88:**9663–9667.
11. **Collins, P. L., R. A. Olmsted, M. K. Spriggs, P. R. Johnson, and A. Buckler-White.** 1987. Gene overlap and attenuation of transcription of the viral polymerase L gene of human respiratory syncytial virus. Proc. Natl. Acad. Sci. USA **84:**5134–5138.
12. **Collins, P. L., and G. W. Wertz.** 1983. cDNA cloning and transcriptional mapping of nine polyadenylated RNAs encoded by the genome of human respiratory syncytial virus. Proc. Natl. Acad. Sci. USA **80:**3208–3212.
13. **Dickens, L. E., P. L. Collins, and G. W. Wertz.** 1984. Transcriptional mapping of human respiratory syncytial virus. J. Virol. **52:**364–369.
14. **Garcin, D., and D. Kolakofsky.** 1992. Tacaribe arenavirus RNA synthesis in vitro is primer dependent and suggests an unusual model of the initiation of genome replication. J. Virol. **66:**1370–1376.
15. **Grosfeld, H., M. G. Hill, and P. L. Collins.** 1995. RNA replication by respiratory syncytial virus (RSV) is directed by the N, P and L proteins: transcription also occurs under these conditions but requires RSV superinfection for efficient synthesis of full-length mRNA. J. Virol. **69:**5677–5686.
16. **Iverson, L. E., and J. K. Rose.** 1981. Localized attenuation and discontinuous synthesis during vesicular stomatitis virus transcription. Cell **23:**477–484.
17. **Johnson, P. R., and P. L. Collins.** 1988. The A and B subgroups of human respiratory syncytial virus: comparison of intergenic and gene-overlap sequences. J. Gen. Virol. **69:**2901–2906.
18. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **54:**367–382.
19. **Kuo, L., R. Fearns, and P. L. Collins.** 1996. The structurally diverse intergenic regions of respiratory syncytial virus do not modulate sequential transcription by a dicistronic minigenome. J. Virol. **70:**6143–6150.
20. **Kuo, L., H. Grosfeld, J. Cristina, M. G. Hill, and P. L. Collins.** 1996. Effect of mutations in the gene-start and gene-end sequence motifs on transcription of monocistronic and dicistronic minigenomes of respiratory syncytial virus. J. Virol. **70:**6892–6901.
21. **Lamb, R. A., and D. Kolakofsky.** 1995. *Paramyxoviridae*: the viruses and their replication, p. 1177–1204. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), Virology, 3rd ed. Raven Press, New York, N.Y.
22. **Samal, S. K., and P. L. Collins.** 1996. RNA replication by a respiratory syncytial 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.
23. **Wagner, R. R., and J. K. Rose.** 1995. *Rhabdoviridae* and their replication. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), Virology, 3rd ed. Raven Press, New York, N.Y.
24. **Yu, Q., R. W. Hardy, and G. W. Wertz.** 1995. Functional cDNA clones of the human respiratory syncytial (RS) virus N, P, and L proteins support replication of RS virus genomic RNA analogs and define minimal *trans*-acting requirements for RNA replication. J. Virol. **69:**2412–2419.
25. **Zamora, M., and S. K. Samal.** 1992. Gene junction sequences of bovine respiratory syncytial virus. Virus Res. **24:**115–121.