Molecular Cloning and Characterization of the RNA Packaging-Defective Retrovirus SE21Q1b

DAWN J. ANDERSON,¹ PEGGY LEE,^{1,2,3} KATHRYN L. LEVINE,¹ JUNSHENG SANG,⁴ SAMIR A. SHAH,⁴
OTTO O. YANG,² PETER R. SHANK,⁴ AND MAXINE L. LINIAL^{1,3*}

Division of Basic Sciences, Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, Washington 98104-2092^{1}; Division of Biology and Medicine, Brown University, Providence, Rhode Island 02912⁴; and Departments of Oral Medicine² and Microbiology,³ University of Washington, Seattle, Washington 98195*

Received 1 July 1991/Accepted 2 October 1991

The nonconditional RNA packaging mutant SE21Q1b contains *cis*- and *trans*-acting defects which cause cellular mRNA, rather than viral genomic RNA, to be nonspecifically packaged into SE21Q1b viral particles. Using genomic libraries of the c-SE21Q1b quail cell line, we have been able to construct a molecular clone of the SE21Q1b provirus. Upon transfection into primary quail embryo fibroblasts, the SE21Q1b molecular clone is able to recapitulate the nonspecific RNA packaging phenotype of the c-SE21Q1b cell line. The RNA packaging phenotypes displayed by several SE21Q1b/avian sarcoma-leukemia virus hybrid provirus constructs have further indicated that sequences responsible for the altered RNA packaging phenotype of SE21Q1b are localized in the left third of the SE21Q1b proviral genome. DNA sequence analysis of this region has revealed that the 5' SE21Q1b deletion has removed 179 bp from the SE21Q1b left long terminal repeat and leader regions. Several differences were detected at the carboxyl terminus of the deduced SE21Q1b nucleocapsid protein sequence in comparison with that of Rous sarcoma virus PR-C. Results of site-directed oligonucleotide mutagenesis experiments indicate, however, that the presence of these residues in the nucleocapsid protein alone is not responsible for the decreased RNA packaging specificity of SE21Q1b.

The specific recognition and packaging of genomic RNA into the viral capsid are important steps in the retrovirus life cycle. The ability of retroviruses to precisely select and encapsidate two genomic RNA molecules amidst a high background of cellular RNA suggests the existence of unique retroviral sequences which mediate this process. *cis*-acting sequences, referred to as ψ (for packaging) sequences, have been identified and characterized (25, 48). These sequences, present at the 5' end of the retroviral genome in the leader (*L*) region, have been found to be required for efficient RNA packaging (43). While additional sequences have been implicated in encapsidation of the retroviral genome (7, 43), ψ is believed to be primarily responsible for specific packaging of genomic RNA (1). The exact nature of the *cis*-acting signal has not been determined, but RNA secondary structure is believed to play a pivotal role.

The presence of a *cis*-acting RNA encapsidation sequence in the retroviral genome suggests the existence of a *trans*-acting factor which specifically recognizes this signal. An obvious candidate for such a factor would be one (or more) of the *gag* proteins, as viral particles containing genomic viral RNA can be produced in the absence of *pol* and *env* gene products (reviewed in reference 8). One of the *gag*-encoded proteins, the nucleocapsid (NC) protein, appears to be required for viral RNA encapsidation (29, 32). A number of mutations in the NC protein coding sequence are lethal for viral replication and specifically have been shown to adversely affect packaging of viral genomic RNA (11, 13, 30-32). Exactly what is required for specific recognition and packaging of genomic retroviral RNA remains a mystery.

A number of retroviral RNA packaging mutants have been isolated and have proven to be useful tools to study the

process of genomic RNA encapsidation. The first such mutant identified was a spontaneous, nonconditional Rous sarcoma virus (RSV) packaging mutant and was designated SE21Q1b (25). Viral particles isolated from the quail cell line containing the SE21Q1b provirus (hereafter referred to as c-SE21Q1b) were found to be composed of apparently normal viral structural proteins and to contain reverse transcriptase. These viral particles were produced in numbers equivalent to those produced by wild-type RSV, but the SE21Q1b viral particles were found to be noninfectious. It was determined that these particles essentially lacked viral RNA, as less than 1% of the encapsidated RNA was identified as virus specific. Instead, SE21Q1b particles were found to have randomly encapsidated cellular mRNA, roughly in proportion to their intracellular concentration (12). Superinfection of c-SE21Q1b with avian leukosis virus (ALV) was not sufficient to relieve the block to packaging of SE21Q1b genomic RNA, suggesting that the SE21Q1b provirus contains a *cis*-acting RNA packaging defect. This defect was believed to be related to a small deletion detected in the SE21Q1b provirus between the left long terminal repeat (LTR) and the *gag* coding sequence. Furthermore, packaging of cellular RNA appeared to be a *trans*-dominant phenomenon; cellular mRNA packaging continues at high levels even in the presence of a superinfecting wild-type virus (25).

We report here the molecular cloning of the SE21Q1b provirus. We have found that this SE21Q1b molecular clone is able to recapitulate the SE21Q1b phenotype upon stable transfection into quail embryo fibroblast (QEF) cells. We have constructed a number of SE21Q1b/avian sarcoma-leukemia virus (ASLV) hybrid proviruses and have used them to determine that only sequences contained in the left third of the SE21Q1b proviral genome are necessary to reproduce the SE21Q1b nonspecific RNA packaging phenotype. We have sequenced the cloned SE21Q1b provirus from

* Corresponding author.

the U3 region of the left LTR through the entire *gag* coding region, allowing us to more precisely characterize the previously identified 5' deletion in the SE21Q1b proviral genome. We have also analyzed the SE21Q1b *gag* protein coding region by comparing it with other avian retroviral *gag* sequences, including those of RSV PR-C, *ev-1*, *ev-2*, and Rous-associated virus-0 (RAV-0) (recently sequenced by us and presented in this report). A number of amino acid differences relative to RSV PR-C were detected in the carboxyl terminus of the deduced SE21Q1b NC amino acid sequence, including one unique to SE21Q1b. Site-directed oligonucleotide mutagenesis experiments indicated, however, that these residues are not alone responsible for the nonspecific RNA packaging phenotype of SE21Q1b.

MATERIALS AND METHODS

DNA manipulations. Plasmid DNA was isolated from *Escherichia coli* by the rapid screening method of Holmes and Quigley (18). Large-scale plasmid isolation was carried out by the method of Ish-Horowitz and Burke (20), with two cycles of cesium chloride-ethidium bromide density gradient centrifugation. All restriction enzymes were purchased from Boehringer Mannheim Biochemicals (Indianapolis, Ind.), Bethesda Research Laboratories (Gaithersburg, Md.), or New England Biolabs, Inc. (Beverly, Mass.) and were used under conditions suggested by the manufacturer. Calf intestinal alkaline phosphatase and T4 DNA ligase were purchased from Boehringer Mannheim Biochemicals. Transformation of *E. coli* with plasmid DNA was performed according to the procedure of Hanahan (14).

c-SE21Q1b genomic library construction. To isolate a molecular clone of the SE21Q1b provirus, a *Sau3A1*-partial genomic library of the c-SE21Q1b quail cell line was prepared. Briefly, genomic DNA was extracted from actively growing c-SE21Q1b quail cells as described by Hughes et al. (19) and partially digested with *Sau3A1*. The DNA was then size fractionated by sucrose gradient density centrifugation (28), ligated with EMBL3 DNA, and packaged *in vitro*, using a two-step λ phage packaging system (17). A second, independent c-SE21Q1b genomic library was similarly constructed by using EMBL4 DNA and Gigapack λ phage packaging extracts (Vector Cloning Systems, San Diego, Calif.). Recombinant phage plaques were screened for the presence of proviral sequence by Southern blot hybridization using ^{32}P -labelled RSV *R* and *gag* region probes. SE21Q1b-containing phage clones were characterized by restriction enzyme analysis and Southern blot hybridization.

Reconstruction of the SE21Q1b provirus. As none of the SE21Q1b molecular phage clones isolated from either library contained the entire provirus, a number of subclones were used to reconstruct an intact SE21Q1b provirus. A 3.0-kb *HindIII*-*SacI* restriction fragment from phage clone ϕSE15 was subcloned into pBluescript II (Stratagene, San Diego, Calif.), yielding construct pRICJ. The 2.1-kb *EcoRI* fragment from ϕKL3700 was subcloned into pVZ1 (16), yielding subclone pRIC. The left third of the SE21Q1b provirus was reconstructed by ligating the 1.7-kb *ApaI*-*EcoRI* fragment from pRICJ with the 1.4-kb *EcoRI*-*SacI* fragment of pRIC, yielding construct pRIJC. The central region of the SE21Q1b provirus was reconstructed from phage clones ϕSE2 and ϕKL3700 . The 3.8-kb *EcoRI* fragment from ϕSE2 was subcloned into pBluescript II to yield subclone pRIA. The 0.66-kb *SacI*-*EcoRI* fragment from subclone pRIC and the 2.7-kb *EcoRI*-*KpnI* fragment of pRIA were ligated together. The 3.36-kb *SacI*-*KpnI* fragment from the resulting construct

was then ligated into *SacI*-, *KpnI*-digested pRIJC, yielding a subclone designated pRIJA, which contained the left half of the SE21Q1b provirus (through the unique *KpnI* site in *pol*) and 5' flanking quail sequences. To obtain the right half of the SE21Q1b provirus, a 7.5-kb *BamHI* fragment from ϕSE2 was subcloned into pBluescript II to yield subclone pBamA. A 6.7-kb *Clal*-*KpnI* fragment (containing the left half of the SE21Q1b provirus) was isolated from pRIJA and ligated with pBamA to reconstruct the complete SE21Q1b provirus with flanking cellular sequences.

Construction of the 5' SE21Q1b/3' RSV hybrid provirus construct pDA2100. A 3.3-kb *Clal*-*HpaI* fragment (containing SE21Q1b left LTR, ΔL , *gag*, and 5' *pol* sequences as well as 5' flanking quail genomic DNA sequences) was isolated. This fragment was ligated into the 11-kb *HpaI*-*Clal* fragment of pJD100 (49) which contained vector-linked RSV PR-A 3' *pol*, *env*, *src*, and right LTR provirus sequences to yield construct pDA2100.

Construction of the hybrid 5' ALV/3' SE21Q1b provirus construct pSE21WT. A 2.3-kb *EcoRI* fragment, which contained the 3' half of the left LTR, leader, and most of *gag*, was isolated from the a molecular clone of the wild-type ALV recombinant NY203 (35, 36) and ligated into subclone pRIJA, replacing the corresponding 2.1-kb SE21Q1b *EcoRI* fragment. The 6.9-kb *Clal*-*KpnI* fragment was then ligated into a slightly modified pBamA to yield construct pSE21WT.

DNA sequence analysis. DNA sequence analysis of the SE21Q1b provirus (from the *EcoRI* site in the left LTR through *gag*) was performed in both directions on overlapping restriction fragment subclones. DNA sequence analysis of RAV-0 *gag* sequences was performed in both directions, using pRAV-0 (45) (provided by K. Conklin, University of Minnesota). DNA sequencing was performed on denatured duplex DNA, using the dideoxy-chain termination method (37) with the Sequenase sequencing system (U.S. Biochemical Corp., Cleveland, Ohio) under conditions suggested by the manufacturer. Sequencing reactions were primed with M13 forward or reverse primers supplied in the sequencing kit or with various oligonucleotide primers synthesized at Brown University on a BioSearch 8600 DNA synthesizer or at the Fred Hutchinson Cancer Research Center Biotechnology Facility on an Applied Biosystems DNA synthesizer. [^{35}S]dATP (1,300 Ci/mmol) was purchased from New England Nuclear Research Products, Boston, Mass.

Site-directed oligonucleotide mutagenesis. Site-directed oligonucleotide mutagenesis of the RSV PR-C NC coding sequence of the pATV8R (provided by E. Hunter, University of Alabama, Birmingham) provirus was performed as described by Kunkel (24). A 2.4-kb *AatII*-*EcoRV* fragment containing the wild-type NC coding sequence of the pATV8R provirus was subcloned into pGEM-5Zf(+) (Promega Corp., Madison, Wis.). An oligonucleotide 45-mer (5'-GGCCTCTCTTCGGGGTCGTGGCCCCGTCTCTGAGCAGCCTGCCGTC-3') was used to introduce the following SE21Q1b NC amino acid changes into the subcloned ATV8R NC coding region: Pro-697 to Ser-697, Gly-700 to Val-700, Pro-701 to Ser-701, and Pro-703 to Gln-703. Introduction of the appropriate SE21Q1b base changes in the pATV8R provirus NC coding sequence was confirmed by DNA sequence analysis. The 2.4-kb *AatII*-*EcoRV* fragment containing the described SE21Q1b NC mutations was then subcloned back into pATV8R, replacing the analogous wild-type sequences.

Cell culture and DNA transfection. QEF cells were obtained from fertile Japanese quail eggs supplied by the animal facility at the Fred Hutchinson Cancer Research

Center. Chicken embryo fibroblast (CEF) cells were obtained from fertile chicken eggs purchased from Highline International, Dallas Center, Iowa. Cell culture maintenance has been previously described (25). Bovine calf and chicken sera were purchased from Hyclone Laboratories, Inc., Logan, Utah, and GIBCO Laboratories, Grand Island, N.Y., respectively. DNA transfection of secondary QEF or CEF cells was performed as described by Chen and Okayama (4).

Soft agar focus assay. The focus-forming abilities of viral particles produced from QEF or CEF cells transfected with the various proviral constructs were assayed essentially as described by Vogt (46) except that Polybrene was used during the initial infection medium at a final concentration of 2 μ g/ml.

Preparation of viral RNA. Virus was collected and purified as described by Linial et al. (25). The procedure used to purify virion RNA has been previously described (1). All solutions (except ethanol) used in the preparation of viral RNA were treated with diethylpyrocarbonate (0.1%) and autoclaved prior to use.

Denaturing PAGE. Core proteins of sucrose-purified virions were analyzed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) as described by Dreyfuss et al. (9).

Preparation of cellular RNA. Total cell RNA was prepared from actively growing transfected cell cultures by the acid guanidinium-phenol-chloroform method of Chomczynski and Sacchi (5).

Virion RNA analysis. Purified virion RNAs were analyzed by standard Northern (RNA) blot hybridization techniques. Virion RNAs were quantitated relative to amounts of viral capsid (CA) protein as estimated from Coomassie-stained SDS-polyacrylamide gels. RNAs were cross-linked to nitrocellulose with UV light ($\lambda = 254$ nm, 120,000 μ J), using a Stratalinker UV cross-linker (Stratagene), and were hybridized with either a 32 P-labelled *gag*- or β -actin-specific probe. Labelled probe was removed from the RNA blot by repeated immersion in a boiling 0.05 $\times$ SSC (1 $\times$ SSC is 0.15 M NaCl plus 0.015 M sodium citrate)-0.01 M EDTA-0.1% (wt/vol) SDS bath. Following immersion of the blot, the heat source was removed, and the blot was allowed to incubate for 30 min. This procedure was repeated as necessary.

Quantitation of encapsidated viral genomic RNA. Whole-band video densitometry was used to determine relative levels of viral genomic RNA packaged into purified viral particles. Densitometric analysis was performed by using the Visage 2000 system (Bioimages/Millipore, Ann Arbor, Mich.).

Nucleotide sequence accession numbers. The SE21Q1b nucleotide sequence data will appear in the EMBL, GenBank, and DDBJ nucleotide sequence data bases under accession number M73496. The pRAV-0 nucleotide sequence data will appear under accession number M73497.

RESULTS

Molecular cloning of the SE21Q1b provirus. It has been hypothesized that the altered RNA packaging specificity of SE21Q1b arises from a defect (or defects) in the provirus rather than from a unique cellular defect. Testing this hypothesis required the molecular cloning of the SE21Q1b provirus, and two independent c-SE21Q1b genomic libraries were constructed as a first step in achieving this end. From one library, 11 SE21Q1b-containing recombinant phages were identified (data not shown). Restriction enzyme characterization of the isolated recombinant phage clones indi-

cated that while none of the recombinants contained the entire SE21Q1b provirus, two recombinant clones, ϕ SE2 and ϕ SE15, did represent the complete provirus (Fig. 1). From a second c-SE21Q1b library, a single SE21Q1b-containing recombinant was identified (data not shown). This recombinant phage clone, designated ϕ KL3700, was found to contain the 5' third of the SE21Q1b proviral genome (Fig. 1).

As none of the recombinant phage clones isolated from either c-SE21Q1b library contained the entire SE21Q1b provirus, it was necessary to reconstruct an intact proviral genome. Several SE21Q1b proviral fragments were subcloned from ϕ SE2, ϕ SE15, and ϕ KL3700 to facilitate this reconstruction. The entire SE21Q1b provirus and flanking quail genomic DNA was represented by four subclones: pRICJ, pRIA, pBamA, and pRIC (Fig. 1). A detailed description of SE21Q1b subclone construction is provided in Materials and Methods. Basically, 5' sequences of the SE21Q1b left LTR as well as 5' flanking quail genomic DNA were obtained from subclone pRICJ. SE21Q1b sequences obtained from pRIC (left LTR and *gag* sequences up to the *SacI* site) and from pRIA (3' *gag* and *pol* sequences) were ligated with those of pRICJ to generate the reconstruction intermediate pRIJA. The right half of the SE21Q1b provirus (3' *pol*, *env*, *src*, and right LTR) and the 3' flanking quail genomic DNA was supplied by subclone pBamA. These SE21Q1b sequences were ligated to those of pRIJA to yield a complete molecular clone of the SE21Q1b provirus, which we have designated pSE21Q (Fig. 1).

pSE21Q sequence analysis. Previous characterization of the SE21Q1b provirus reported by two of us had revealed that the SE21Q1b provirus contained a deletion estimated to be 150 bp in length and that this deletion was located 300 to 600 bp from the left (*gag-pol*) end of the provirus (41). The availability of an SE21Q1b molecular clone has allowed us to more fully characterize the SE21Q1b provirus by DNA sequence analysis. We chose to concentrate our efforts on the left end of the SE21Q1b provirus, specifically from the *EcoRI* restriction site in the U3 region of the left LTR through the entire *gag* region. Our rationale behind the selection of these sequences was twofold. First, sequencing of this region would allow us to more precisely determine the size and location of the previously identified SE21Q1b 5' deletion. Secondly, it would allow us to characterize the coding sequences of the SE21Q1b *gag* proteins, one or more of these proteins having been previously implicated as *trans*-acting viral RNA packaging factors.

The sense-strand nucleotide sequence of the left third of the SE21Q1b provirus, from the *EcoRI* restriction site in U3 through the *gag* protein coding region, is shown in Fig. 2. To illustrate a number of its features, the SE21Q1b nucleotide sequence is presented in relation to the nucleotide sequences from corresponding regions of several other avian retroviruses. The selection of these particular sequences was based on previous characterization of SE21Q1b as a recombinant virus (25). The SE21Q1b parent virus PR-E 21 was originally obtained as a subgroup E recombinant following infection of *ev-1*⁺ *ev-2*⁺ chicken cells with RSV PR-B (26). The *ev-1* locus is an endogenous provirus which, although expressed at low levels, does not produce any infectious virus (3, 15). The *ev-2* locus has been shown to be the endogenous provirus which encodes infectious RAV-0 (2). We determined the nucleotide sequence shown for RAV-0, using the RAV-0 molecular clone pRAV-0.

Alignment of the SE21Q1b and RSV PR-C sequences has allowed us to determine that the 5' deletion in the SE21Q1b

FIG. 1. Restriction maps of SE21Q1b molecular clone pSE21Q and of c-SE21Q1b phage clones and subclones used in reconstruction of the SE21Q1b provirus. Hatched bars, SE21Q1b proviral sequences; open bars, flanking quail genomic sequences.

provirus genome is in fact a deletion of 179 nucleotides (from nucleotides 96 to 274, inclusive) (Fig. 2). Sequences lost from the SE21Q1b genome as a result of this deletion include the right (3') half of the U5 inverted repeat, the entire positive-strand primer binding site, and over half of the *L* region.

SE21Q1b is, as previously described, a subgroup E recombinant virus, and denaturing PAGE of labelled SE21Q1b virus core proteins had suggested that at least one of its *gag* proteins, CA protein p27₀, was inherited originally from RAV-0 (25). CA p27₀ has been shown to be directly related to CA p27 despite its retarded electrophoretic mobility (39). A comparison of the SE21Q1b sequence, at both the nucleotide and deduced amino acid levels (Fig. 2 and 3, respectively), with corresponding sequences of RAV-0 supports the earlier characterization of SE21Q1b as a RAV-0 recombinant (25). The nucleotide sequences of the SE21Q1b and RAV-0 CA protein coding region are identical except for three nucleotide differences. One nucleotide difference at position 1305 means that the SE21Q1b CA protein (like the RSV PR-C CA) contains a methionine residue at amino acid position 309 whereas the RAV-0 CA contains a threonine residue. The other two nucleotide differences (nucleotides 1200 and 1282) represent conservative nucleotide changes, meaning that no differences exist between the CA proteins of SE21Q1b and RAV-0 (or, in fact, RSV PR-C) at the deduced

amino acid level. It is interesting to note that SE21Q1b and RSV PR-C contain the same nucleotides at these positions.

The SE21Q1b and RAV-0 p10 coding regions are identical, and the NC and protease (PR) protein coding regions each differ by only a single nucleotide (Fig. 2). The nucleotide difference between SE21Q1b and RAV-0 NC coding sequences is located at position 2088. The T at this position in the SE21Q1b sequence would lead to an encoded serine residue, whereas the C in RAV-0 would lead to proline (Fig. 3). The nucleotide difference in the PR coding region is located at nucleotide 2340 (Fig. 2). This difference means the RAV-0 PR protein would contain a proline at amino acid residue 654 whereas the SE21Q1b protease would contain an arginine (Fig. 3). Interestingly, the RSV PR-C protease also contains an arginine residue at this location, although the triplet encoding it is not the same as the one encoding Arg-654 in the SE21Q1b protease. The PR proteins of SE21Q1b and RAV-0 both differ from that of RSV PR-C at amino acid residue 659, where SE21Q1b and RAV-0 encode a valine whereas RSV PR-C encodes a leucine (Fig. 3).

Earlier SDS-PAGE analysis of SE21Q1b viral core proteins had revealed that the SE21Q1b matrix (MA) protein is present in two forms, p19 α and p19 β , both of which are related to the RSV PR-C and RAV-0 p19 protein (25, 39). Other recombinant viruses have also been shown to possess these alternative forms of the MA p19 protein (39, 40).

FIG. 2. Nucleotide sequences of the leader and *gag* regions of SE21Q1b and RAV-0. The nucleotide sequence of the sense strand of the SE21Q1b provirus molecular clone is depicted on the top line. The second line depicts the nucleotide sequence of the RAV-0 provirus molecular clone pRAV-0. Lines 3 and 4 represent the nucleotide sequences of the *ev-2* and *ev-1* loci, respectively. Sequences for these two loci were obtained from reference 47. Line 5 depicts the nucleotide sequence of RSV PR-C and was obtained from reference 38. A dash indicates sequence identity, and an asterisk indicates no nucleotide. Assignment of nucleotide numbers is based on the conventional numbering for RSV PR-C (38). PBS, primer binding site.

SE21Q1b	1.	CAGCAGGGCC	AGGGGGGTGA	CACACCTCGG	GGGGCGGAAC	AGCCAAGGGC	GGAGCCAGGG	CACGCGGGTC	TGGCCCCCTGG	GCCGGCCCTG	ACTGACTGGG	1046
pRAV-0	2.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
RSV PRC	3.	-----	-----	-----C-	-----	-----T-----	-----	-----T-----	A-----T-----	-----	-----	
ev-2	4.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
ev-1	5.	-----	-----	-----T	-----	-----	-----	-----	-----	-----	-----	
p10 — CA												
SE21Q1b	1.	CAAGGATCAG	GGAGGAGCTT	GCGAGTACAG	GTCCGCCCGT	GGTGGCCATG	CCTGTAGTGA	TTAAGACAGA	GGGACCCGCG	TGGACCCCTC	TGGAGCCAAA	1146
pRAV-0	2.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
RSV PRC	3.	-----G-----	-----	-----T-----	-----	-----	-----	-----	-----	-----	-----	
ev-2	4.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
ev-1	5.	-----	-----	-----T-----	-----	-----	-----	-----	-----	-----	-----	
SE21Q1b	1.	ATTGATCACA	AGACTGGCTG	ATACGGTCAG	GACCAAGGGC	TTACGATCCC	CGATCACTAT	GGCAGAAGTG	GAAGCGCTCA	TGTCCTCCCC	GCTGCTGCCG	1246
pRAV-0	2.	-----	-----	-----	-----	-----	-----C-----	-----	-----	-----	-----	
RSV PRC	3.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
SE21Q1b	1.	CATGACGTCA	CGAATCTAAT	GAGAGTTATT	TTAGGGCCCTG	CCCCATATGC	CTTATGGATG	GACGCTTGGG	GAGTCCAAC	ACAGACGGTT	ATAGCGGCAG	1346
pRAV-0	2.	-----	-----	-----	-----A-----	-----	-----C-----	-----	-----	-----	-----	
RSV PRC	3.	-----	-----	-----	-----	-----	-----	-----	-----	C-----A-----	-----	
SE21Q1b	1.	CCACTCGCGA	CCCCCGACAC	CCAGCGAACG	GTCAGGGCGG	GGGGGAACGG	ACTAACTTGG	ATCGTTTAAA	GGGTTTGGCG	GATGGAATGG	CCGGCAATCC	1446
pRAV-0	2.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
RSV PRC	3.	-----	-----	-----	-----	-----	-----T---A-----	-----C-----	-----C---A---T-----	-----	TG-----T-----	
SE21Q1b	1.	AGAGGGTCAG	GCTGCATTAT	TAAGACCGGG	GGAAGTGGTT	GCTATTACGG	CGTCGGCTCT	CCAGGCGTTT	AGAGAGGTCG	CTCGGTTGGC	GGAAACCACA	1546
pRAV-0	2.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
RSV PRC	3.	-C-----	-----	-----	-----T-----	-----	-----	-----	-----T-----	C---C-----	-----TG	
SE21Q1b	1.	GACCCGTGGG	CGGAAATTAC	GCAGGGACCA	TCTGAGTCTT	TTGTGGATTT	TGCCAATCGT	CTTATAAAGG	CGGTTGAAGG	GTCAGATCTC	CCACCTTCCG	1646
pRAV-0	2.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
RSV PRC	3.	-GT---A----	---C---C-T-----	-----	-----	-----T-----	-----G-----	-----	-----G-----	-----	-G-----	
SE21Q1b	1.	CGCGAGCTCC	GGTGATCATT	GACTGCTTTA	GGCAGAAGTC	ACAGCCAGAT	ATTCAGCAGC	TTATACGGGC	AGCACCCCTC	ACGCTGACCA	CCCCAGGAGA	1746
pRAV-0	2.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
RSV PRC	3.	---G-----	-----	-----	-----	-----	-----	-----A-----	-----	-----	-----	
SE21Q1b	1.	GATAATCAAA	TATGTGCTAG	ACAGGCAGAA	GACTGCCCTT	CTTACGGATC	AAGGCATAGC	CGCGGCCATG	TCGCTCTGCTA	TTCAGCCCTT	AGTTATGGCA	1846
pRAV-0	2.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
RSV PRC	3.	-----	-----	-----	-----	-----	-----	-----	-----	C-----	A-----	
SE21Q1b	1.	GTAGTCAATA	GAGAGAGGGA	TGGACAAACT	GGGTCGGGTG	GTCGTGCCCG	AGGGCTCTGC	TAACTTGTG	GATCCCCGGG	ACATTATCAG	GCGCAGTGCC	1946
pRAV-0	2.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
RSV PRC	3.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
SE21Q1b	1.	CGAAAAAACG	AAAGTCAGGA	AACAGCCGTG	AGCGATGTCA	GCTGTGTGAC	GGGATGGGAC	ACAACGCTAA	ACAGTGACGA	AGGCGGGATG	GCAACCAGGG	2046
pRAV-0	2.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
RSV PRC	3.	-----	G-----	-----	-----	-T-----A--	-----	-----	-----T-G	-A-----	-----	
SE21Q1b	1.	CCAACGCCCA	GGAAGAGGCC	TCTCTTCGGG	GTCGTGGCCC	GTCTCTGAGC	AGCCTGCCGT	CTCGTTAGCG	ATGACAATGG	AACATAAAGA	TGCCCCCTTG	2146
pRAV-0	2.	-----	-----	-----	-----C-----	-----	-----	-----	-----	-----	-----	
RSV PRC	3.	-----	---A---T-----	-----	-C-----	-G-C---CC-	-----	-----	-----	-----	-----	

FIG. 2—Continued.

Conklin et al. have demonstrated that a p19 α and a p19 β protein are encoded by endogenous virus *ev-1* sequences (6). These researchers suggested that recombinant viruses expressing p19 α and p19 β had, in all likelihood, acquired their MA protein coding region via recombination with *ev-1*

sequences. A comparison of the SE21Q1b MA nucleotide and deduced amino acid sequences with those of the *ev-1* locus suggests that SE21Q1b has indeed inherited at least part of its MA protein coding sequence from *ev-1*. Specifically, the 3' end of the SE21Q1b MA coding sequence as well

SE21Q1b	1.	GTTAGGGTCA	TTCTGACTAA	CACTGGGAGT	CATCCGGTCA	AACAGCGTTC	GGTGTATATC	ACCGCGCTGT	TGGACTCTGG	AGCGGACATC	ACTATTATTT	2246
pRAV-0	2.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
RSV PRC	3.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	

SE21Q1b	1.	CAGAGGAGGA	CTGGCCACCC	GATTGGCCAG	TGATGGAGGC	CGCGAACCCG	CAGATCCATG	GGATAGGAGG	GGGAATTCCC	ATGCGAAAAAT	CCCGGGATAT	2346
pRAV-0	2.	-----	-----	-----	-----	-----	-----	-----	-----	-----	---C---	
RSV PRC	3.	-----	T-----	-----	-----	-----	-----	-----	-----	-----	-T--T--G--	

SE21Q1b	1.	GATAGAGGTG	GGGGTTATTA	ACCGAGACGG	GTCTTTGGAG	CGACCCCTGC	TCCTCTTCCC	CGCAGTAGCT	ATGGTTAGAG	GGAGTATCCT	AGGAAGAGAT	2446
pRAV-0	2.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
RSV PRC	3.	-----	T-----	-----	-----	-----	-----	-----	-----	-----	-----	

SE21Q1b	1.	TGCTGTCAGG	GCCTAGGGCT	CCGCTTGACA	AATTTATAGG	GAGGGCCACT	GTTCCTCACTG	TTGCGCTACA	TCTGGCTATT	CCGCTCAAAT	GGAAGCCAGA	2546
pRAV-0	2.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
RSV PRC	3.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	

SE21Q1b	1.	CCACACGCCT	GT	2558
pRAV-0	2.	-----	-----	
RSV PRC	3.	-----	-----	

FIG. 2—Continued.

as the spacer region between MA and p10 appear to have originated with *ev-1* (Fig. 2 and 3). Unlike RAV-0 or RSV PR-C, SE21Q1b and *ev-1* encode an alanine, lysine, and methionine at positions 133, 143, and 144, respectively, in the MA protein as well as a glutamine and threonine at positions 156 and 161, respectively, in the spacer region. It is the lysine residue present at position 143 which is believed to be responsible for the increased mobility of p19 β relative to p19 (47). Whether SE21Q1b has acquired the rest of its MA coding region from *ev-1* or from RAV-0 cannot be conclusively determined. A comparison of the MA nucleotide sequence of *ev-1* and RAV-0/*ev-2* reveals that they are virtually identical except for three nucleotide differences at the 3' end (47). We should, however, note that an A-to-C transversion relative to *ev-2* was detected at nucleotide 623 in the RAV-0 molecular clone that we sequenced (Fig. 2). This change would lead to the substitution of glutamine for lysine at residue 72 in the MA protein (Fig. 3). We believe that this difference is a feature unique to this particular isolate of RAV-0.

Construction of SE21Q1b/RSV hybrid proviruses. It has generally been assumed, though never conclusively demonstrated, that the observed altered RNA packaging specificity of SE21Q1b can be attributed to SE21Q1b proviral sequences rather than to a unique c-SE21Q1b cellular defect. By obtaining an SE21Q1b molecular provirus clone, it became possible to examine this hypothesis directly. Furthermore, as both the leader and *gag* regions have been shown to play important roles in packaging of retroviral genomic RNA, we were particularly interested in determining whether these sequences in SE21Q1b are responsible for its decreased RNA packaging specificity.

To identify the sequences in the SE21Q1b genome responsible for the altered RNA packaging phenotype, we constructed two recombinant proviral molecular clones. One of our hybrid proviruses, designated pDA2100, was composed of 5' sequences of SE21Q1b (left LTR, Δ 179 L, *gag*, and 5' *pol*) fused to 3' sequences of wild-type RSV PR-A (3' *pol*, *env*, *src*, and right LTR) (Fig. 4). Alternatively, the second hybrid provirus, designated pSE21WT, was composed basically of 5' ALV and 3' SE21Q1b sequences. In this construct, the pSE21Q Δ L region and *gag* through most of the

PR coding region were replaced with the corresponding sequences from the wild-type ALV recombinant NY203, a RAV-60 class recombinant virus which is believed to have acquired most of its sequence from a RAV-1 parent (Fig. 4) (35, 36, 42).

Transfection of SE21Q1b constructs. Each of our proviral constructs, pSE21Q, pDA2100, and pSE21WT, was transfected into secondary QEF cells. Morphological differences between the individual transfected cultures were clearly apparent early in passaging. Shortly after the first passage, QEFs transfected with pSE21WT (the 5' ALV/3' SE21Q1b hybrid provirus) acquired a transformed phenotype which was virtually identical to that displayed by QEFs transfected with either an RSV PR-A or an RSV PR-C molecular proviral clone (data not shown). Rapid transformation of the SE21WT culture suggested the presence of an infectious virus. Infectivity of SE21WT viral particles was further examined by assaying supernatants collected from the SE21WT-transformed culture for focus-forming ability. As shown in Table 1, SE21WT is able to transform QEFs with an efficiency comparable to that of RSV PR-C (ATV8R) and RSV PR-A (JD100).

In contrast, QEF cultures transfected with pSE21Q (the SE21Q1b molecular clone) and pDA2100 (the 5' SE21Q1b/3' RSV PR-A proviral hybrid) did not begin to display a transformed phenotype until at least passage 4, and even then not all cells appeared to be transformed. This slow and incomplete transformation of the SE21Q and DA2100 cultures suggested to us that a infectious transforming virus was not being spread through either culture, rather that increased multiplication of cells stably transfected with either pSE21Q or pDA2100 was eventually responsible for the appearance of the cultures' transformed phenotype. The ability of SE21Q and DA2100 viral particles to mediate a productive viral infection was examined by assaying mass culture supernatants for focus-forming ability. Transformation of QEFs by SE21Q and DA2100 viral supernatants occurred at a level only slightly above background, with essentially the same efficiency as SE21Q1b viral supernatants (Table 1). Viral particles produced by the SE21Q and DA2100 cultures thus appear to be incapable of mediating a productive

	1	MA	
SE21Q1b	1.	MEAVIKVISSACKTYCGKTSPSKKEIGAMLSLLQKEGLLMSPSDLYSPGS	
pRAV-0	2.	-----	
RSV PRC	3.	-----	
ev-2	4.	-----	
ev-1	5.	-----	
	51		
SE21Q1b	1.	WDPIITAALSQRAMVLGKSGELKKTWGLVLGALKAAAREEQVTSEQAKFWLGL	
pRAV-0	2.	-----	
RSV PRC	3.	-----	
ev-2	4.	-----	
ev-1	5.	-----	
	101		
SE21Q1b	1.	GGGRVSPPGPECIEKPATERRIDKGEEVGETTAAQRDAKMAPEKMATPKTV	
pRAV-0	2.	-----	
RSV PRC	3.	-----	
ev-2	4.	-----	
ev-1	5.	-----	
	151	MA	
SE21Q1b	1.	GTSCYQCGTATGCNCATASAPPPYVGSGLYP SLAGVGEQQGQGDDTPRG	p10
pRAV-0	2.	-----	
RSV PRC	3.	-----	
ev-2	4.	-----	
ev-1	5.	-----	
	201		
SE21Q1b	1.	AEQPRAEPGHAGLAPGPALTDWARIREELASTGPPVVAMPVVV IKTEGPAW	p10 CA
pRAV-0	2.	-----	
RSV PRC	3.	-----	
ev-2	4.	-----	
ev-1	5.	-----	
	251		
SE21Q1b	1.	TPLEPKLITRLADTVRTKGLRSPITMAEVEALMSSPLLPHDVTNLMRVIL	
pRAV-0	2.	-----	
RSV PRC	3.	-----	
	301		
SE21Q1b	1.	GPAPYALWMDAWGVQLQTVIAAATRDRHPANGQGRGERTNLDRLKGLAE	
pRAV-0	2.	-----	
RSV PRC	3.	-----	
	351		
SE21Q1b	1.	GMAGNPEGQAALLRPGELVAITASALQAFREVARLAEPDTPWAEITQGFS	
pRAV-0	2.	-----	
RSV PRC	3.	-----	
	401		
SE21Q1b	1.	ESFVDFANRLIKAVEGSDLPSPSARAPVIIDCFRQKSQPDIIQQLIRAAPST	
pRAV-0	2.	-----	
RSV PRC	3.	-----	
	451		
SE21Q1b	1.	LTTPGEIIKYVLDRQKTAPLTDQGI AAAMSSAIQPLVM AVVNRRERDGTG	CA NC
pRAV-0	2.	-----	
RSV PRC	3.	-----	
	501		
SE21Q1b	1.	SGGRARGLCYTCS PGHYQAQCPKKRKSGNSRERCQLCDGMGHNAKQCRR	
pRAV-0	2.	-----	
RSV PRC	3.	-----	
	551		
SE21Q1b	1.	RDGNQGRPGRGLSSGSWPVSEQPAVSLAMTMEHKDRPLVRVILTNTGSH	NC PR
pRAV-0	2.	-----	
RSV PRC	3.	-----	
	601		
SE21Q1b	1.	PVKQRSVYITALLD SGADITII SEEDWPTDWPVME AANPQIHGIGGGIPM	
pRAV-0	2.	-----	
RSV PRC	3.	-----	
	651		
SE21Q1b	1.	RKSRDMI EVGVINRDGSLERPLLLFP AVAMVRGSI LGRDCLQGLGLRLTNL	
pRAV-0	2.	-----	
RSV PRC	3.	-----	

FIG. 3. Deduced amino acid sequence of the *gag* protein coding region of SE21Q1b and RAV-0. Line 1 depicts the predicted sequence of the SE21Q1b Pr76^{gag}, beginning with the first amino acid residue of the MA (p19) protein. Line 2 represents the predicted sequence of the pRAV-0 Pr76^{gag}. A dashed line indicates sequence identity. Lines 3 and 4 depict the deduced amino acid sequences for endogenous virus ev-2 (RAV-0) and ev-1, respectively. Line 5 represents the predicted amino acid sequence of RSV PR-C Pr76^{gag}.

FIG. 4. Diagrams of the *SE21Q1b* molecular clone pSE21Q and of the two *SE21Q1b*/ALSV hybrid provirus molecular clones pDA2100 and pSE21WT. Hatched bars, *SE21Q1b* proviral sequences; open bars, flanking quail genomic sequences; stippled bars, NY203 (ALV) sequences; diagonally striped bars, ATV8R (RSV PR-C) sequences; vertically striped bars, JD100 (RSV PR-A) sequences.

infection. This, of course, is one of the hallmarks of viral particles produced by the *SE21Q1b* quail cell line.

Northern hybridization analysis of SE21Q, DA2100, and SE21WT virion RNA. *SE21Q1b* had previously been shown to package RNA into viral particles with decreased specificity. Rather than specifically packaging its own genomic RNA, *SE21Q1b* was found instead to randomly package

cellular mRNAs (12). To determine the specificity of RNA packaging of virus derived from the *SE21Q1b* molecular clone and the *SE21Q1b*/ALSV hybrid provirus constructs, viral particles were harvested from *SE21Q*, *DA2100*, and *SE21WT* mass cultures. Culture supernatant was collected, and virion RNA was isolated from purified viral particles and characterized by Northern hybridization analysis.

TABLE 1. Focus-forming ability of RSV virions^a

Cells	Focus-forming units/ml							Mock
	SE21Q1b	SE21Q	DA2100	SE21WT	PL45NC	ATV8R	JD100	
QEF	<0.1	2.8	3.8	8.1×10^5	ND	ND	7.5×10^4	<0.1
CEF	<0.1	ND	ND	ND	1.4×10^7	1.5×10^7	ND	<0.1

^a Twenty-four-hour culture supernatants were harvested from fully infected QEF cells and clarified of cell debris by low-speed centrifugation. Aliquots were filtered on secondary CEF or QEF cells as previously described (46). Foci were counted 10 to 12 days after infection, ND, not done.

FIG. 5. Northern hybridization analysis of virion RNAs probed with *gag* sequences. Viral supernatants from primary QEF cells transfected with various SE21Q1b constructs and wild-type RSV were harvested and purified, and virion RNA was extracted as described in Materials and Methods. Virion RNA samples were electrophoresed through a formaldehyde agarose gel, blotted to nitrocellulose, and hybridized with a 32 P-labelled *gag* probe. Equivalent amounts of virion RNA, based on relative amounts of viral capsid protein, were electrophoresed in each lane. Lanes: 1, c-SE21Q1b cellular RNA; 2, SE21Q1b; 3, SE21Q; 4, DA2100; 5, SE21WT; 6, PL45NC; 7, ATV8R (RSV PR-C); 8, RSV PR-A (JD100); 9, mock (pBR322); 10, ALV control. Hybridization was allowed to proceed overnight at 42°C in 50% formamide. Blots were washed at 65°C in 2× SSC (0.3 M NaCl, 0.03 M sodium citrate)–0.1% SDS and in 0.1× SSC–0.1% SDS. Arrow and number denote molecular weight (in thousands).

To detect encapsidation of genomic viral RNA, the virion RNA blot was first probed with a 32 P-labelled *gag*-specific probe. As can be seen in Fig. 5, a small yet detectable amount of viral RNA was found to be packaged into SE21Q viral particles (lane 3). A similar amount of viral RNA was also found packaged into DA2100 viral particles (lane 4). SE21Q1b, SE21Q, and DA2100 appear to be significantly less efficient than RSV PR-C (lane 7) or RSV PR-A (lane 8) at packaging their genomic RNA. Roughly 30-fold less viral RNA is detected in the SE21Q1b, SE21Q, and DA2100 virion RNA samples relative to the amount detected in either of the wild-type RSV samples. Thus, the same decreased efficiency of viral RNA packaging observed for SE21Q1b is also displayed by the SE21Q1b molecular clone SE21Q and by the 5' SE21Q1b/3' RSV hybrid DA2100. Furthermore, as the SE21Q1b, SE21Q, and DA2100 proviral genomes are all believed to lack a functional viral RNA packaging signal due to the deletion in the leader region, it seems likely that the viral RNA detected in the respective virions was packaged by a nonspecific mechanism.

In contrast, SE21WT appears to be able to efficiently package its genomic RNA (Fig. 5, lane 5). While slightly less viral RNA appears to be packaged by SE21WT than by RSV PR-C (ATV8R) and RSV PR-A (JD100) (compare lane 5 with lanes 7 and 8, respectively), other Northern hybridizations that we have performed with the same RNA samples suggest that this slight difference is most probably due to sample loading variations rather than to less efficient encapsidation of viral RNA by SE21WT.

To determine whether SE21Q and DA2100, like SE21Q1b, nonspecifically package cellular mRNA, we rehybridized the virion RNA blot with a 32 P-labelled chicken β -actin probe. We chose to probe for packaged β -actin message because the efficiency of cellular mRNA packaging by SE21Q1b

FIG. 6. Northern hybridization analysis of virion RNAs probed with β -actin sequences. The virion RNA blot used in the Northern analysis shown in Fig. 5 was stripped of labelled *gag* probe as described in Materials and Methods. The RNA blot was reprobed with a 32 P-labelled chicken β -actin probe. Hybridization and wash conditions have been previously described. Lanes: 1, c-SE21Q1b cellular RNA; 2, SE21Q1b; 3, SE21Q; 4, DA2100; 5, SE21WT; 6, PL45NC; 7, ATV8R (RSV PR-C); 8, JD100 (RSV PR-A); 9, mock (pBR322); 10, ALV control. Arrow and number denote molecular weight (in thousands).

appears to correlate directly with the intracellular concentration of the message (12). It seemed likely that the mRNA of a cellular structural protein would be present at a concentration sufficient to allow its relatively frequent packaging, consequently making it easier to detect.

The results of this Northern hybridization are shown in Fig. 6. Like SE21Q1b virions (lane 2), SE21Q and DA2100 viral particles have encapsidated detectable amounts of β -actin mRNA (lanes 3 and 4, respectively). The intensities of the actin-specific bands suggest that while SE21Q1b and SE21Q package actin mRNA with near equal efficiency, DA2100 appears to package actin message with greater frequency. The near-equivalent amounts of viral RNA packaged by SE21Q1b, SE21Q, and DA2100, as judged by the uniform intensities of the virus-specific bands (Fig. 5, lanes 2, 3, and 4, respectively), suggest that the observed variation in actin-specific band intensities is not simply a sample loading artifact. We have, however, not noticed a consistent increase in β -actin mRNA encapsidation by DA2100, relative to SE21Q1b or SE21Q, in other Northern analyses.

In contrast, we could not detect any β -actin message encapsidated into SE21WT viral particles (Fig. 6, lane 5). RSV PR-C (ATV8R) (lane 7) or RSV PR-A (JD100) (lane 8) virions also do not appear to have encapsidated significant amounts of β -actin message, although, upon overexposure of this Northern blot, we did notice a small amount of β -actin mRNA packaged into RSV PR-C (ATV8R) viral particles (data not shown). This, however, is not unexpected since low-level packaging of cellular RNA (1 to 2% of the total RNA encapsidated) by RSV has previously been observed (1, 25). The fact that β -actin mRNA was not encapsidated by SE21WT particles while significant amounts of this message were detected in DA2100 virions again suggests that sequences present in the SE21Q1b *gag* gene play a role in altering the specificity of RNA packaging.

Oligonucleotide mutagenesis of the RSV PR-C NC coding sequence. As the sequences responsible for the altered RNA packaging specificity of SE21Q1b had been localized to the 5' end of the proviral genome, we were interested in characterizing the nature of the SE21Q1b *trans*-acting packaging

defect. While the 179-bp deletion at the 5' end of the *SE21Q1b* proviral genome undoubtedly contributes to the inefficient packaging of *SE21Q1b* genomic RNA, it, in all likelihood, is not responsible for *SE21Q1b*'s random packaging of cellular RNA. Other proviruses which contain similar 5' deletions have not been found to nonspecifically package RNA (21, 33, 44). As previously discussed, several *gag* proteins have been implicated at putative *trans*-acting RNA packaging factors, and at least one, NC, has been shown to bind RNA, albeit nonspecifically (29). Interestingly, the one feature of the *gag* protein coding sequence completely unique to *SE21Q1b* is the C-to-T transition at nucleotide 2088, which leads to a proline-to-serine substitution at the carboxyl terminus of the deduced *SE21Q1b* NC amino acid sequence (Fig. 2 and 3). Just downstream of this serine residue, the *SE21Q1b* NC protein contains a cluster of three amino acids (a valine, a serine, and a glutamine at positions 670, 671, and 673, respectively) which differ from those in the RSV PR-C NC protein (Fig. 3). These residues are encoded by sequences presumably inherited from RAV-0. If the NC protein (independently or as part of the *gag* precursor polyprotein Pr76^{gag}) indeed mediates specific packaging of viral RNA, we reasoned that the unique *SE21Q1b* Ser-667 residue in conjunction with Val-670, Ser-671, and Gln-673 might be responsible for the relaxed RNA packaging specificity of *SE21Q1b*.

To examine this hypothesis, we used an oligonucleotide identical to the last 45 nucleotides of the *SE21Q1b* NC coding sequence to mutagenize the NC coding sequence of RSV PR-C ATV8R, leading to substitution of the Ser-667, Val-670, Ser-671, and Gln-673 residues of the *SE21Q1b* NC for the Pro-667, Gly-670, Pro-671, and Pro-673 residues of the ATV8R protein. The mutagenized proviral construct, designated pPL45NC, was then transfected into secondary QEFs. Shortly after the first passage, the PL45NC cell culture acquired a transformed phenotype (data not shown). Transformation of the PL45NC culture progressed at a rate essentially equivalent to that of the control ATV8R culture, suggesting the presence of an infectious transforming virus. Infectivity of PL45NC viral particles was examined by assaying PL45NC culture supernatants for the ability to form foci in a soft agar assay. As shown in Table 1, PL45NC is able to transform CEFs with an efficiency comparable to that of RSV PR-C ATV8R. To determine the nature of the RNA packaged by PL45NC, PL45NC virion RNA was extracted from purified viral particles and subjected to Northern hybridization analysis with ³²P-labelled *gag* and β -actin probes. On the basis of the intensities of the *gag*-specific bands, PL45NC (Fig. 5, lane 6) appears to package viral RNA approximately 50-fold more efficiently than does *SE21Q1b* (lane 2), and as well as its wild-type ATV8R parent (lane 7). We could not detect a β -actin-specific band (lane 6), suggesting that little if any β -actin mRNA is encapsidated into PL45NC viral particles. PL45NC thus appears to have retained the RNA packaging specificity of wild-type RSV.

DISCUSSION

Information obtained from the studies presented here have allowed us to more fully characterize the nonconditional RNA packaging mutant *SE21Q1b*. Initial characterizations had suggested that the *SE21Q1b* provirus contained both *cis*- and *trans*-acting RNA packaging defects (25). Using a number of partial *SE21Q1b* molecular clones obtained from two genomic libraries of the c-*SE21Q1b* cell line, we have been able to reconstruct a complete *SE21Q1b* molecular proviral

clone, pSE21Q (Fig. 1). Analysis of the transforming ability of the virus encoded by the *SE21Q1b* molecular clone and of the RNA encapsidated into *SE21Q* virions has shown that the nonspecific RNA packaging phenotype of *SE21Q1b* can be directly ascribed to the *SE21Q1b* provirus rather than to a cellular factor. A nonspecific RNA packaging phenotype was also displayed by the 5' *SE21Q1b*/3' RSV hybrid DA2100, one of two recombinant proviruses that we constructed. In contrast, the 5' ALV/3' *SE21Q1b* hybrid provirus *SE21WT* displayed an RNA packaging phenotype essentially equivalent to that of wild-type RSV. The fact that DA2100 displays a phenotype essentially identical to that of *SE21Q1b* while that of *SE21WT* resembles wild-type RSV suggests that sequences responsible for the altered RNA packaging phenotype of *SE21Q1b* can be localized to *SE21Q1b* 5' and *gag* proviral sequences.

We have further characterized *SE21Q1b* by sequencing most of the 5' end of the *SE21Q1b* provirus, from the *EcoRI* site in U3 of the left LTR through the entire *gag* region (Fig. 2). Alignment of the *SE21Q1b* nucleotide sequence with the corresponding RSV PR-C sequence has allowed us to determine that the 5' deletion previously detected in the *SE21Q1b* proviral genome is actually 179 bp in length and results in the loss of the 3' half of the U5 inverted repeat, the positive-strand primer binding site, and a large portion of the *L* region. We believe that the deleted *L*-region sequences underlie the dramatically decreased packaging of the *SE21Q1b* genomic RNA by eliminating, truncating, or otherwise interfering with the *cis*-acting viral RNA packaging signal ψ . Without a functional packaging signal, viral genomic RNA cannot be efficiently encapsidated, leading to decreased virus infectivity. We believe that other aspects of the 179-bp deletion contribute to the noninfectious nature of *SE21Q1b* viral particles. As RSV genomic RNA is present in significant amounts in an infected cell, one would expect that even random packaging of *SE21Q1b* genomic RNA would be sufficient to allow at least a small number of infectious virus particles to be produced. However, the absence of a positive-strand primer binding site would severely restrict reverse transcription, making replication of *SE21Q1b* genomic RNA highly unlikely. Furthermore, *SE21Q1b*'s lack of the critical 3' half of the U5 inverted repeat would preclude normal integration of any viral DNA intermediate.

Like *SE21Q1b*, other RSV packaging mutants contain deletions in their *L* sequences. Another naturally occurring packaging mutant TK15 contains a 237-bp deletion in its provirus which has removed *L* sequences just 3' of the positive-strand primer binding site (23, 33). The efficiency with which TK15 genomic RNA is packaged is significantly reduced compared with that of its parental virus (22). But unlike *SE21Q1b*, TK15 is able to produce a low titer of infectious virus particles, presumably related to the presence of an intact primer binding site in the TK15 genome (22). An engineered packaging-defective RSV mutant, RAV-1 ψ^- , contains a 153-bp *L*-region deletion which has removed proviral sequences from the 3' half of the primer binding site to the *SacI* site (44). RAV-1 ψ^- is not only unable to package its own genomic RNA, but, like *SE21Q1b*, is also replication defective. A more detailed analysis of RSV *L*-region sequences has demonstrated that a 30-bp sequence just upstream of the *SacI* site appears to be essential for packaging of genomic viral RNA (21). This sequence is absent from the genomes of *SE21Q1b*, TK15, and RAV-1 ψ^- . It has, however, recently been shown that deletion of the *L* region is not alone sufficient to yield particles which contain randomly packaged cellular RNAs (1).

Although the mutation(s) leading to random packaging of cellular RNAs maps to the *gag* region, we have not yet determined precisely what features of the SE21Q1b *gag* gene are responsible. The only aspect of the SE21Q1b *gag* sequence which appeared to be unique to the SE21Q1b genome was the T at nucleotide 2088, leading to substitution of serine for proline in the carboxyl terminus of the SE21Q1b deduced NC amino acid sequence (Fig. 2 and 3). This predicted change in the SE21Q1b NC protein was particularly intriguing given the role that the NC protein appears to play in viral RNA encapsidation. However, when substituted for the corresponding RSV PR-C NC amino acid in conjunction with three RAV-0-derived amino acids (Val-670, Ser-671, and Gln-673), we could not detect nonspecific packaging of β -actin mRNA. We have yet to examine the effect of these amino acid changes in a genome which also has the SE21Q1b 179-bp *L*-region deletion.

The only mutations in NC shown to decrease RNA packaging specificity have so far been localized to residues of the single Cys-His motif of the Moloney murine leukemia virus NC protein (13). Other studies have shown that deletion of either of the two RSV NC Cys-His repeats or point mutations in the proximal repeat leads to a decrease in the efficiency of viral RNA binding and encapsidation (10, 31, 32). The SE21Q1b NC, like other avian retroviral NC proteins, contains two degenerate Cys-His repeats which are identical to those of the RSV PR-C NC save for one residue. The SE21Q1b NC contains an aspartate at position 539 in the distal repeat, whereas RSV PR-C contains an asparagine. The only other amino acid differences detected between the SE21Q1b and RSV PR-C NC proteins, other than those that we have already tested, are at positions 550 and 561, where the SE21Q1b NC contains arginine residues while the RSV PR-C NC contains lysine residues (Fig. 3). It is interesting that except for the serine residue at position 667 (which does not appear to be responsible for the SE21Q1b RNA packaging phenotype), the deduced amino acid sequence of the SE21Q1b NC protein is identical to that of RAV-0. If NC is responsible for mediating SE21Q1b's decreased specificity of RNA packaging, one might expect RAV-0 to display a similar phenotype. We are now examining the specificity of RNA encapsidation by RAV-0.

A number of recent publications suggest that specific viral RNA packaging is mediated not by an individual *gag* proteins but by the unprocessed *gag* precursor polyprotein (34). In fact, sequence-specific binding of the human immunodeficiency virus type 1 *gag* polyprotein Pr55^{gag} to human immunodeficiency virus type 1 RNA has recently been reported (27). These results imply that other domains of the *gag* polyprotein, in addition to NC, may participate in determining RNA packaging specificity. The actual *trans*-acting region may consist of juxtaposed polyprotein domains, brought together by specific folding of the *gag* polyprotein. It is possible that the decrease in RNA packaging specificity of SE21Q1b may stem from altered folding of Pr76^{gag}, perhaps attributable to the recombinant nature of the SE21Q1b *gag* coding region. We should note that while SE21Q1b does package RNA nonspecifically, it has not entirely lost the ability to recognize the viral RNA packaging signal ψ (1). One might therefore expect only a subtle change in the SE21Q1b *trans*-acting RNA packaging factor. It seems possible that a slight alteration in the conformation of the *gag* precursor polyprotein, mediated by one or more of the amino acids encoded by sequences inherited from RAV-0 or *ev-1*, might be responsible for the decreased RNA packaging specificity of SE21Q1b. Studies currently in progress are

aimed at determining the exact nature of the *trans*-acting defect responsible for the decreased RNA packaging specificity of SE21Q1b.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grants CA 18282 and AI 27291 to M.L.L. and CA 32980 to P.R.S. D.J.A. was supported by Public Health Service postdoctoral fellowship CA 08944-01 from the National Cancer Institute. P.L. was supported by dental scientist training grant DE 00161.

P.R.S. thanks Siegrid Reinsch for expert technical assistance.

REFERENCES

1. Aronoff, R., and M. L. Linial. 1991. Specificity of retroviral RNA packaging. *J. Virol.* 65:71-80.
2. Astrin, S., L. B. Crittenden, and E. G. Buss. 1980. *Ev-2*, a genetic locus containing structural genes for endogenous virus, codes for Rous-associated virus type 0 produced by line 7(2) chickens. *J. Virol.* 33:250-255.
3. Baker, B., H. L. Robinson, H. E. Varmus, and J. M. Bishop. 1981. Analyses of endogenous avian retrovirus DNA and RNA: viral and cellular determinants of retrovirus gene expression. *Virology* 114:8-22.
4. Chen, C., and H. Okayama. 1987. High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* 7:2745-2752.
5. Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium-thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156-159.
6. Conklin, K. F., J. M. Coffin, H. L. Robinson, M. Groudine, and R. Eisenman. 1982. Role of methylation in the induced and spontaneous expression of the avian endogenous virus *ev-1*: DNA structure and gene products. *Mol. Cell. Biol.* 2:638-652.
7. DeGudicibus, S. J., B. Gentile, R. S. Bhatt, M. S. Poonian, and D. W. Stacey. 1986. Studies of retroviral packaging, p. 59-65. *In* J. E. Celis, A. Graessmann, and A. Loyter (ed.), *Microinjection and organelle transplantation techniques*. Academic Press, London.
8. Dickson, C., R. N. Eisenman, H. Fan, E. Hunter, and N. Teich. 1984. Protein biosynthesis and assembly, p. 514-648. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), *RNA tumor viruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
9. Dreyfuss, G., S. A. Adam, and Y. O. Choi. 1984. Physical change in cytoplasmic messenger ribonucleoproteins in cells treated with inhibitors of mRNA transcription. *Mol. Cell. Biol.* 4:415-423.
10. Dupraz, P., S. Oertle, C. Meric, P. Damay, and P. F. Spahr. 1990. Point mutations in the proximal Cys-His box of Rous sarcoma virus nucleocapsid protein. *J. Virol.* 64:4978-4987.
11. Fu, X., R. A. Katz, A. M. Skalka, and J. Leis. 1988. Site-directed mutagenesis of the avian retrovirus nucleocapsid protein, pp12. *J. Biol. Chem.* 263:2140-2145.
12. Gallis, B., M. Linial, and R. N. Eisenman. 1979. An avian oncovirus mutant deficient in genomic RNA: characterization of the packaged RNA as cellular messenger RNA. *Virology* 94:146-161.
13. Gorelick, R. J., L. E. Henderson, J. P. Hanser, and A. Rein. 1989. Point mutants of Moloney murine leukemia virus that fail to package viral RNA: evidence for specific RNA recognition by a "zinc finger-like" protein sequence. *Proc. Natl. Acad. Sci. USA* 85:8420-8424.
14. Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166:557-580.
15. Hayward, W. S., S. B. Braverman, and S. M. Astrin. 1980. Transcriptional products and DNA structure of endogenous avian proviruses. *Cold Spring Harbor Symp. Quant. Biol.* 44:1111-1121.
16. Henikoff, S., and M. Eghtedarzadeh. 1987. Conserved arrangement of nested genes at the *Drosophila Gart* locus. *Genetics* 117:711-725.
17. Hohn, B., and K. Murray. 1977. Packaging recombinant DNA

- molecules into bacteriophage particles in vitro. *Proc. Natl. Acad. Sci. USA* 74:3259-3263.
18. Holmes, D. S., and M. Quigley. 1981. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* 114:193-197.
 19. Hughes, S. H., P. R. Shank, D. H. Spector, H.-J. Kung, J. M. Bishop, H. E. Varmus, P. K. Vogt, and M. L. Breitman. 1978. Proviruses of avian sarcoma virus are terminally redundant, coextensive with unintegrated linear DNA and integrated at many sites in rat cell DNA. *Cell* 15:1397-1410.
 20. Ish-Horowicz, D., and J. F. Burke. 1981. Rapid and efficient cosmid vector cloning. *Nucleic Acids Res.* 9:2989-2998.
 21. Katz, R. A., R. W. Terry, and A. M. Skalka. 1986. A conserved *cis*-acting sequence in the 5' leader of avian sarcoma virus RNA is required for packaging. *J. Virol.* 59:163-167.
 22. Kawai, S., and T. Koyama. 1984. Characterization of a Rous sarcoma virus mutant defective in packaging its own genomic RNA: biological properties of mutant TK15 and mutant-induced transformants. *J. Virol.* 51:147-153.
 23. Koyama, T., F. Harada, and S. Kawai. 1984. Characterization of a Rous sarcoma virus mutant defective in packaging its own genomic RNA: biochemical properties of mutant TK15 and mutant-induced transformants. *J. Virol.* 51:154-162.
 24. Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* 82:488-492.
 25. Linial, M., E. Medeiros, and W. S. Hayward. 1978. An avian oncovirus mutant (SE21Q1b) deficient in genomic RNA: biological and biochemical characterization. *Cell* 15:1371-1381.
 26. Linial, M., and P. E. Neiman. 1976. Infection of chick cells by subgroup E viruses. *Virology* 73:508-520.
 27. Luban, J., and S. P. Goff. 1991. Binding of human immunodeficiency virus type 1 (HIV-1) RNA to recombinant HIV-1 *gag* polyprotein. *J. Virol.* 65:3203-3212.
 28. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 29. Meric, C., J. L. Darlix, and P.-F. Spahr. 1984. It is Rous sarcoma virus p12 and not p19 that binds tightly to Rous sarcoma virus RNA. *J. Mol. Biol.* 173:531-538.
 30. Meric, C., and S. P. Goff. 1989. Characterization of Moloney murine leukemia virus mutants with single-amino-acid substitutions in the Cys-His box of the nucleocapsid protein. *J. Virol.* 63:1558-1568.
 31. Meric, C., E. Gouilloud, and P.-F. Spahr. 1988. Mutations in Rous sarcoma virus nucleocapsid protein p12 (NC): deletions of Cys-His boxes. *J. Virol.* 62:3328-3333.
 32. Meric, C., and P.-F. Spahr. 1986. Rous sarcoma virus nucleic acid-binding protein p12 is necessary for viral 70S RNA dimer formation and packaging. *J. Virol.* 60:450-459.
 33. Nishizawa, M., T. Koyama, and S. Kawai. 1985. Unusual features of the leader sequence of Rous sarcoma virus packaging mutant TK15. *J. Virol.* 55:881-885.
 34. Oertle, S., and P. F. Spahr. 1990. Role of the *gag* polyprotein precursor in packaging and maturation of Rous sarcoma virus genomic RNA. *J. Virol.* 64:5757-5763.
 35. Rettenmier, C. W., and H. Hanafusa. 1977. Structural protein markers in the avian oncovirus. *Virology* 24:850-864.
 36. Robinson, H. L., B. M. Blais, P. N. Tschlis, and J. M. Coffin. 1982. At least two regions of the viral genome determine the oncogenic potential of avian leukosis viruses. *Proc. Natl. Acad. Sci. USA* 79:1225-1229.
 37. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
 38. Schwartz, D. E., R. Tizard, and W. Gilbert. 1983. Nucleotide sequence of Rous sarcoma virus. *Cell* 32:853-869.
 39. Shaikh, R., M. Linial, S. Brown, A. Sen, and R. Eisenman. 1979. Recombinant avian oncoviruses. II. Alterations in the *gag* proteins and evidence for intragenic recombination. *Virology* 92:463-481.
 40. Shaikh, R., M. Linial, J. Coffin, and R. Eisenman. 1976. Recombinant avian oncoviruses. I. Alterations in the precursor to the internal structural proteins. *Virology* 87:326-338.
 41. Shank, P. R., and M. Linial. 1980. Avian oncovirus mutant (SE21Q1b) deficient in genomic RNA: characterization of a deletion in the provirus. *J. Virol.* 36:450-456.
 42. Shank, P. R., R. J. Schatz, J. M. Jensen, P. N. Tschlis, J. M. Coffin, and H. L. Robinson. 1985. Sequences in the *gag*-*pol*-5' *env* region of avian leukosis viruses confer the ability to induce osteopetrosis. *Virology* 145:94-104.
 43. Sorge, J., W. Ricci, and S. H. Hughes. 1983. *cis*-Acting RNA packaging locus in the 115-nucleotide direct repeat of Rous sarcoma virus. *J. Virol.* 48:667-675.
 44. Stoker, A. W., and M. J. Bissell. 1988. Development of avian sarcoma and leukemia virus-based vector-packaging cell lines. *J. Virol.* 62:1008-1015.
 45. Tschlis, P. N., L. Donehower, G. Hager, N. Zeller, R. Malavarca, S. Astrin, and A. M. Skalka. 1982. Sequence comparison in the crossover region of an oncogenic avian retrovirus recombinant and its nononcogenic parent: genetic regions that control growth rate and oncogenic potential. *Mol. Cell. Biol.* 2:1331-1338.
 46. Vogt, P. K. 1969. Focus assay of Rous sarcoma virus, p. 198-211. In K. Habel and N. Saltzman. (ed.), *Fundamental techniques in virology*. Academic Press, New York.
 47. Vogt, V. M., R. B. Pepinsky, and L. E. Southard. 1985. Primary structure of p19 species of avian sarcoma and leukemia viruses. *J. Virol.* 56:31-39.
 48. Watanabe, S., and H. M. Temin. 1982. Encapsidation sequences for spleen necrosis virus, an avian retrovirus, are between the 5' long terminal repeat and the start of the *gag* gene. *Proc. Natl. Acad. Sci. USA* 79:5986-5990.
 49. Wilkerson, V. W., D. L. Bryant, and J. T. Parsons. 1985. Rous sarcoma virus variants that encode *src* proteins with an altered carboxy terminus are defective for cellular transformation. *J. Virol.* 55:314-321.