Nucleic Acids Research, Vol. 18, No. 1 *© 1990 Oxford University Press*

A study of the dimer formation of Rous sarcoma virus RNA and of its effect on viral protein synthesis *in vitro*

Eric Bieth, Caroline Gabus and Jean-Luc Darlix
LaboRetro, Centre de Recherches de Biochimie et de Génétique Cellulaires du CNRS, 118 route de Narbonne, 31062 Toulouse, France

Received September 20, 1989; Revised and Accepted December 1, 1989

ABSTRACT

The genetic material of all retroviruses examined so far is an RNA dimer where two identical RNA subunits are joined at their 5′ ends by a structure named dimer linkage structure (DLS). Since the precise location and structure of the DLS as well as the mechanism and role(s) of RNA dimerization remain unclear, we analysed the dimerization process of Rous sarcoma virus (RSV) RNA. For this purpose we set up an *in vitro* model for RSV RNA dimerization. Using this model RSV RNA was shown to form dimeric molecules and this dimerization process was greatly activated by nucleocapsid protein (NCp12) of RSV. Furthermore, RSV RNA dimerization was performed in the presence of complementary 5′^{32}P-DNA oligomers in order to probe the monomer and dimer forms of RSV RNA. Data indicated that the DLS of RSV RNA probably maps between positions 544 – 564 from the 5′end. In an attempt to define sequences needed for the dimerization of RSV RNA, deletion mutageneses were generated in the 5′ 600 nt. The results showed that the dimer promoting sequences probably are located within positions 208 – 270 and 400 – 600 from the 5′ end and hence possibly encompassing the cis-acting elements needed for the specific encapsidation of RSV genomic RNA. Also it is reported that synthesis of the polyprotein precursor Pr76^{gag} is inhibited upon dimerization of RSV RNA. These results suggest that dimerization and encapsidation of genome length RSV RNA might be linked in the course of virion formation since they appear to be under the control of the same cis elements, E and DLS, and the trans-acting factor nucleocapsid protein NCp12.

INTRODUCTION

Cells infected with Rous Sarcoma Virus (RSV) produce virus specific RNAs that are the 22 S, 28 S, 35 S and 70 S RNAs (1,2 and references herein). In infected cells the 22 S and 28 S RNAs are the messengers for the pp60 v-Src and the Env precursor, respectively, and are found in minute amounts in virions probably because they lack a 5′GAG sequence located between the splice donor site at position 398 and position 550, and more precisely around position 545 (3,4). The 35 S RNA appears to assume two functions : messenger RNA for the synthesis of the protein precursors Pr76^{gag} and Pr180$^{gag\text{-}pol}$, and precursor to the genomic 70 S RNA. A single 70 S RNA is present in mature RSV virions and it consists of an RNA dimer where two identical 35 S RNA of 9312 nucleotides (nt) are linked at their 5′ end by the dimer linkage structure (DLS)(5−7). This structural feature is common to both avian and mammalian type C retroviruses and has been mapped by electron microscopy around positions 480−540 from the 5′ end of RSV-PrA RNA (7). Thus the DLS of RSV RNA dimer appears to be located close to the element necessary for genomic RNA encapsidation present in GAG gene around position 545 (4,8 and references herein). In addition the DLS must involve tertiary nucleic acid structures as indicated by electron microscopic and biochemical data : the 5′ end of the two 35 S RNA appears to be in the same direction in the RSV 70 S RNA dimer and although the point of linkage between the subunits probably consists of about 50 nt, the 70 S RNA dimer is converted to monomeric 35 S RNA upon heating at 60°C (3,7−9).

In the capsid of RSV virions the 70 S RNA interacts with a large number of nucleocapsid protein (NC) molecules and a few reverse transcriptase (RT) molecules (10−12). RSV NCp12, like most other retroviral NC proteins, is derived from the carboxy terminal end of the Pr76^{gag} precursor (8) and it is a small basic protein with nucleic acid binding activity (13−15). In addition it contains two copies of the conserved Cys-His boxes (Cys-Xaa2-Cys-Xaa4-His-Xaa4-Cys) sometimes implied in Zn++ and nucleic acid binding (16−17). Alteration of the Cys-His boxes of RSV and MoMuLV NC by deletions or site directed mutagenesis abolishes the infectivity of these viruses, which also show a marked defect in packaging their genomic 70 S RNA (18−20). Despite their relative small size, retroviral NC proteins have a variety of other functions. For instance it has been reported that they promote the annnealing of the replication primer tRNA to the RNA genome (21). Also preliminary data indicate that RSV NC can direct dimerization of retroviral RNA, a process that could be relevant to reverse transcription, since it has been shown that RNA template switching occurs during the synthesis of proviral DNA (21,22).

Although, the DLS has been visualized by electron microscopy (7), its precise location and the mechanism of RNA dimerization as well as its possible function clearly remained to be identified. In an attempt to characterize parameters of RSV RNA dimerization *in vitro* and sequences involved in the DLS, we initiated a study of the structural and functional features of the

DLS of RSV RNA. SP6-derived RSV RNA having the leader sequence plus the 5′ portion of GAG gene was synthesized and its ability to form dimeric RNA molecules was monitored. We tentatively mapped the DLS of RSV RNA by probing monomeric and dimeric RNA molecules with complementary DNA oligonucleotides, and we tried to define cis-elements needed for RSV RNA dimerization. The possible function of the nucleocapsid protein NCp12 in RSV RNA dimerization was also investigated.

The results reported here show that RSV RNA spontaneously dimerizes, implying direct interactions between the two RNA molecules, and that dimerization is catalysed by the viral NC protein. The sequences necessary for RNA dimerization have been located in two domains : In the leader sequence at positions 208−270 where an encapsidation element named 'E' has been mapped (23), and between positions 400−600 that encompasses a second element required for genomic RNA packaging (4). The DLS probably is located within this second element at positions 544−564. In addition we show that upon dimerization RSV RNA cannot direct synthesis of the viral polyprotein precursor $Pr76^{gag}$.

Thus it is proposed that the dimerization of RSV RNA causes a translational inhibition of genome length RNA and therefore promotes the encapsidation of RSV genomic RNA into virions.

MATERIALS AND METHODS

Plasmid constructions

E.coli strain 1035 (recA−) was used for plasmid manipulation and preparation. For generation of viral RNA by *in vitro* transcription, derivatives of plasmid pSP64 and pSP65 (Pharmacia) were constructed. RSV-PrBtd cloned in Lambda EMBL3 (24) was partially digested with EcoRI (position −52 with respect to the Cap site in the two LTRs) and inserted in pSP65. One plasmid was selected, pLAD4, which contains a 7,2 kb insert with the 5′Leader, the GAG, POL and ENV genes and U3 lacking the 3′ 52nt. pLAD4 was digested with SacI (position +255) and viral fragment was inserted in pSP65 (SacI site); plasmid pLAD5 resulted. The RSV DNA fragment from pLAD5 was further digested at the 5′ end with Bal 31 exonuclease, inserted into pSP64 and two clones of interest were recovered: pLAD6 and pLAD7 with viral 5′ end at positions 372 and 390 respectively. Plasmid pd12 (23)(a gift of Dr. R. Katz) contains a deletion from position 208 to 270 in the 5′ leader of RSV RNA, it was cut with EcoRI and viral fragment from positions −52 to +2312 was inserted in pSP65 and pLAD8 resulted. Plasmid pLAD4 was also cut with BamHI and viral fragment from positions +532 to +1916 was inserted in Blue Scribed (BS) vector and pLAD9 resulted. Plasmid pJS2 (21)(25)(a gift of Dr.M. Stoltzfus) contains the v-SRC cDNA flanked by the LTRs, it was cut with EcoRI and viral fragment inserted into pSP65 and pSPSRC1 resulted. Plasmid pSPSRC1 was cut with SacI and viral fragment inserted into pSP65; plasmid pSPSRC2 resulted. Plasmid DNAs were isolated by alkaline lysis, followed by RNase treatment and phenolisation (26).

Proteins, enzymes and sera

RSV NC (NCp12) and MoMuLV NC (NCp10) were purified from their respective virions by DNA cellulose chromatography (27). RSV capsid (CA) and matrix (MA) proteins as well as protease (PR) were purified by HPLC chromatography and provided by Pr.P.F. Spahr, Geneva. Purified HIV-1 CA and MA proteins, produced in *E.coli*, were kindly provided by Dr. H.V. Kolbe, Transgene, Strasbourg. Purified AMV DNA polymerase, T4 phage gene 32 protein and *E.coli* RecA protein were given to us by M. Gazeau (Geneva), Dr. D. Caput (Toulouse) and Pr. R. Devoret (Orsay), respectively. Proteins were analysed by polyacrylamide gel electrophoresis in SDS (PAGE-SDS) and following coomassie blue staining they appeared to be over 95% pure. Antisera against RSV NC, CA and RT were prepared in rabbits.

Cell culture and virion production

RSV-PrBtd infected chicken embryofibroblasts and MoMuLV-infected NIH 3T3 cells were cultured in DMEM containing 5% FCS at 37°C in the presence of 5% CO2. Virions were purified by two rounds of centrifugation and the amount of virion proteins estimated from coomassie blue staining or immunoblotting following PAGE-SDS electrophoresis (24).

RNA preparation

For *in vitro* synthesis of RSV RNA, 5μg of plasmid DNA were linearised at the XhoI site (+622) for pLAD4, pLAD5 and pLAD8, the EcoRI site (+2312) for pLAD6 and pLAD7, the BglII site (+1630) for pLAD9 and the PvuII site (position +670 with respect to initiator AUG of v-SRC gene) for pSRC.

For *in vitro* synthesis of RSV RNA containing the complete GAG coding sequence, plasmid DNAs were linearised 3′ to GAG at the HindIII site (+2730) for pLAD4, pLAD5, pLAD6 and pLAD8. For *in vitro* synthesis of RSV RNA containing the complete v-SRC gene, pSPSRC was linearised 3′ to v-SRC at the PvuI site (+ 9180 with respect to PrC sequence). The different plasmids were then transcribed with SP6 RNA polymerase, or T7 RNA polymerase for pLAD9, under conditions suggested by the manufacturer (Promega). Following DNase treatment, RNA was phenol extracted, chromatographed over Sephadex G75, ethanol precipitated in the presence of 0,3 M Na-acetate (this was left 6 to 12 hs at −20°C before centrifugation). The RNA precipitate was dissolved in double distilled sterile water and kept at −20°C. Beef liver $tRNA^{Trp}$ was purified as described previously (28). 5′-^{32}P labelled $tRNA^{Trp}$ was prepared using T4 polynucleotide Kinase and ^{32}P-ATP, and the labelled tRNA purified on a 10% urea-polyacrylamide gel. RSV 70S (dimeric genome) and 35S (monomer) RNA were prepared as already described (29).

DNA oligomers

Synthetic DNA oligomers complementary to positions 211−232, 364−383, 520−541 and 544−564 were made using an Applied Biosystem Synthesizer. They were purified by 8% polyacrylamide gel electrophoresis in 7 M urea and 50 mM Tris-Borate pH 8,3. DNA oligomers were 5′-^{32}P labelling with ^{32}P-ATP and T4 Kinase, and specific radioactivity was about 3×10^7 cpm/μg DNA oligomer.

NC assays

These were performed in 10 μl reactions comprising 40 mM Tris-HCl pH 7,0, 50 mM NaCl, 1mM MgCl2, 10 μM ZnCl2, 5 mM DTT, 0,3−0,5 μg *in vitro* generated RNA, 10−100 ng NC or 0,5−1 μg of another viral protein, or *E.coli* protein, for 15 to 30 min at 37°C. The RSV RNA-protein complex was analysed by 0,8% agarose gel electrophoresis in 50 mM Tris-Acetate pH 7,1, 5 mM 2-mercaptoethanol at 3 V/cm for 4 h, with buffer recirculation. Following electrophoresis, gels were stained for

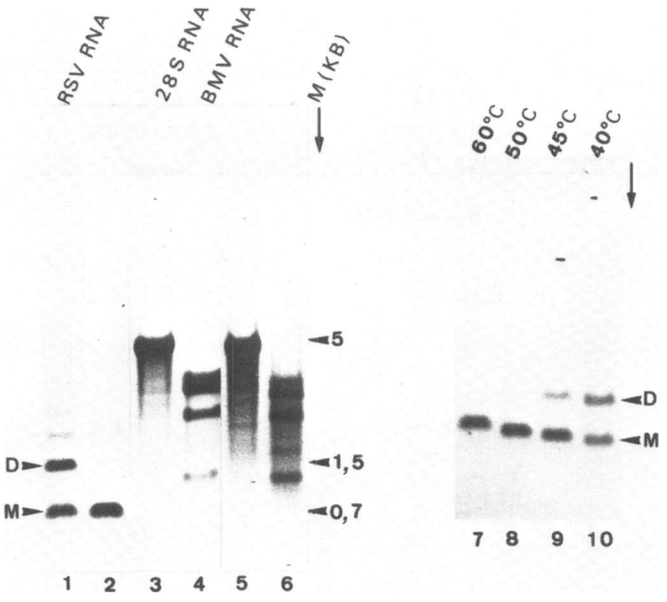

Figure 1: Agarose gel electrophoresis of SP6-generated RSV RNA dimers. Transcription of pLAD4 (RSV) plasmid (cut with XhoI) with SP6 RNA polymerase and RNA purification were carried out as reported in materials and methods. 28 S rRNA was purified from CEF cells and BMV RNA was provided by Amersham.

Lanes 1 to 6 :

1/ RSV RNA native. 2/ RSV RNA denatured 1 min at 95 °C.
3/ 28 S rRNA native. 4/ BMV RNA native.
5/ 28 S rRNA denatured. 6/ BMV RNA denatured.

Note that more than 50% of RSV RNA is in the dimer form, and that trimeric RSV RNA is visible. M corresponds to markers in Kilobases.

Lanes 7 to 10 :

RSV RNA was incubated in dimerization buffer for 5 min at 60°C, 50°C, 45°C and 40°C and loaded during electrophoresis. M & D are monomer and dimer RNA ; the arrow indicates direction of electrophoresis.

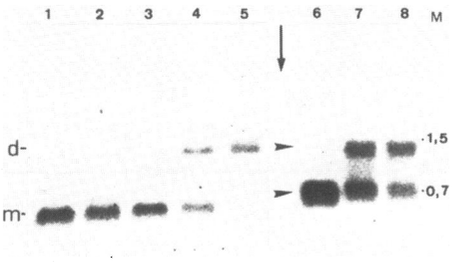

Figure 2: Nucleocapsid protein NCp12 trans-activates RSV RNA dimerization. RSV RNA dimerization was performed in conditions of NC protein assay (see methods) with 0,5 μg of heat denatured RSV RNA (1 min at 95°C). At the end of the incubation RNA was treated with phenol (see methods) and analyzed by agarose gel electrophoresis in native conditions.

Lane 1 : control RSV RNA.
Lane 2 : RSV RNA incubated for 3 hs at 37°C.
Lane 3 : RSV RNA incubated for 15 min at 37°C.
Lane 4 : RSV RNA incubated with 40 ng NCp12 for 15 min at 37°C.
Lane 5 : RSV RNA incubated with 100 ng NCp12 for 15 min at 37°C.
Lane 6 : RSV RNA incubated for 15 min at 37°C.
Lane 7 : RSV RNA incubated with 50 ng NCp10 for 15 min at 37°C.
Lane 8 : RSV RNA incubated with 100 ng NCp10 for 15 min at 37°C.

m and d are the monomer and dimer forms, and arrow is direction of electrophoresis. M indicates RNA markers in kilobases synthesized in vitro and heat denatured before electrophoresis.

10 min with 1 μg/ml ethidium bromide. If only the RNA was to be analysed, following incubation the protein was removed by phenol-SDS extraction and RNA analysed by electrophoresis through a 0,8% agarose gel run in 50 mM Tris-Borate pH 8,3, 1 mM EDTA and at 7 V/cm. Gels were then washed with water and the RNA was visualised by ethidium bromide staining as above. To calculate the ratio of NCp12 to RNA, molecular weight of NCp12 was estimated to be 9 Kda (8, 15).

Protein synthesis in the rabbit reticulocyte lysate (RRL)

RSV RNA with the complete GAG or v-SRC sequence was generated *in vitro* in the presence of a Cap analogue at 0,5 mM (Pharmacia) in order that the 5′ end be capped. RSV RNAs were translated in the RRL at 70% the original concentration with 5−30 μg/ml of RNA, 1 mCi/ml ^{35}S-Methionine and at 31°C for 1 h. ^{35}S-Proteins were analysed by electrophoresis through 8% polyacrylamide gel in 0,1% SDS. Following electrophoresis gels were fixed with 10% acetic acid, 30% methanol, dried and autoradiographed for one day (see also ref. 29).

Determination of RNA secondary structures by computer folding

A modified version of the RNA2 program of Zuker & Stiegler (30) was used. The 5′ RNA sequences of a number of avian retroviruses have been reviewed (31), and in a attempt to determine the secondary structure of the 5′ 600 nt of RSV RNA the known sequences were folded without modifications, or with the following modifications: The primer binding site should be annealed to the primer $tRNA^{Trp}$ (3,21), the GAG initiator AUG as well as the splice donor (SD) are probably non-base paired (32,33). In addition a number of G residues are accessible to T1 RNase and then residues next to these G's should not be base paired (34−36).

RESULTS

Spontaneous dimerization of RSV RNA *in vitro*

In an attempt to more precisely characterize the DLS and the RNA dimerization process, large amounts of SP6 or T7-derived RSV RNA having the leader and the 5′GAG sequences (including the putative DLS around positions 480−540) were synthesized *in vitro* . After synthesis and purification in a low ionic strength buffer (see methods for details) RSV RNA was analysed by agarose gel electrophoresis in native conditions or after heat denaturation (1 min at 100°C).

Native RSV RNA with the complete 380 nt Leader and ending at position 622 gave two major bands (figure 1, lane 1). The lower band had the expected size of 680 nt while the upper one corresponded to a dimer size RNA which disappeared upon heat denaturation (lane 2). Sometimes very small amounts of trimer size RNA were visible and also disappeared upon heat denaturation. RSV RNA dimer could be reformed after denaturation by prolonged (> 6 hs) incubation of RNA in 0,3 M Na acetate pH 7,0 (not shown). This spontaneous dimerization seemed to be specific to RSV RNA since neither 28S rRNA (lane 3) nor Brome mosaic virus (BMV) RNA (lane 4) gave rise to dimeric RNA molecules (lanes 5 & 6 are after 1 min at 100°C).

Finally, electron microscopic examination of retroviral RNA with the leader and 5′GAG synthesized *in vitro* confirmed that both monomeric and dimeric RNA molecules were present (results not shown and see reference 37).

In order to measure the denaturation temperature of RSV RNA dimer, RNA was incubated at 60°C, 50°C, 45°C and 40°C, in

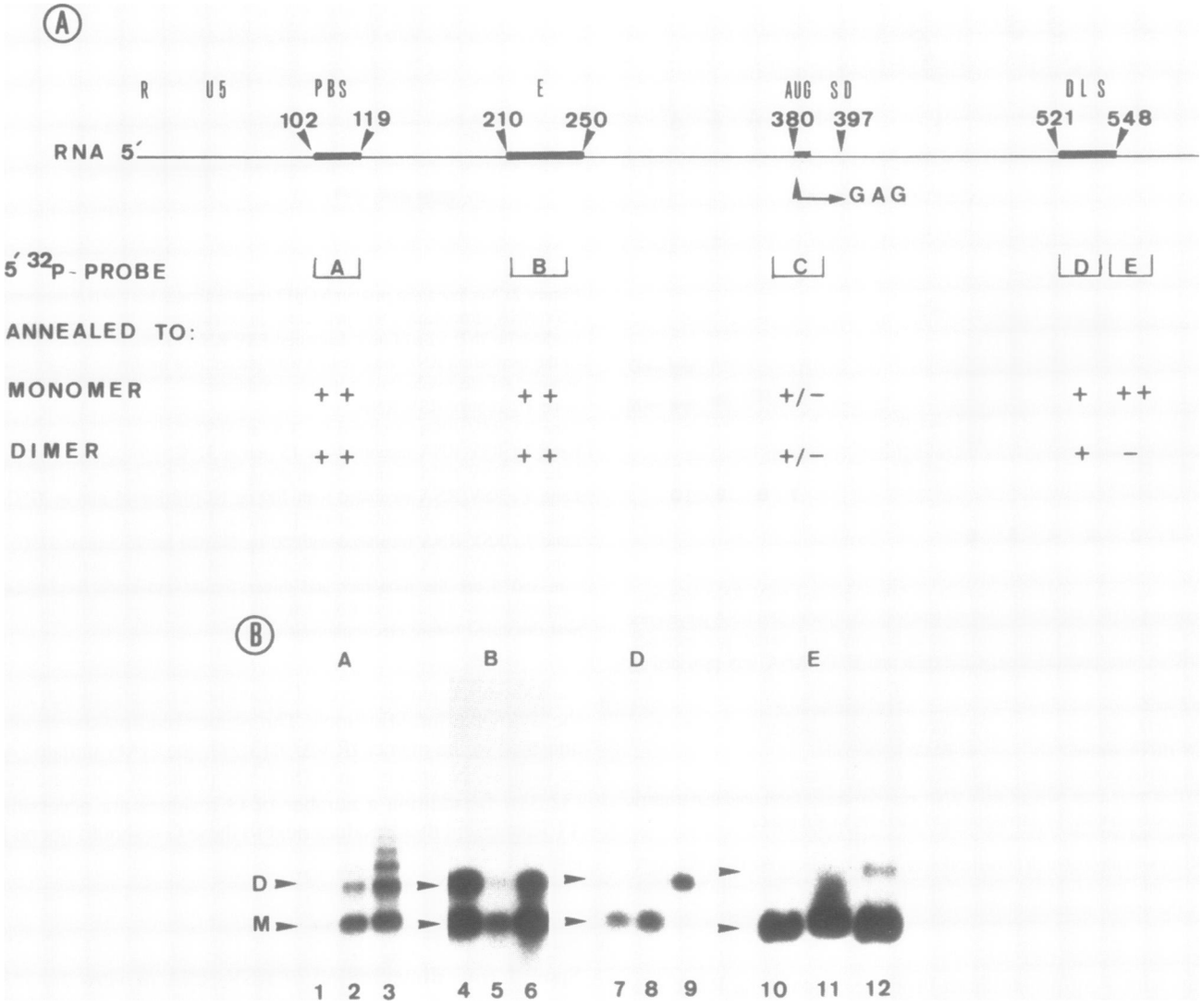

Figure 3: Tentative mapping of the DLS by hybridizing DNA oligomers to RSV RNA.
A) The 5' 600 nt of RSV RNA is drawn. R (terminal repeat), U5, PBS (primer binding site), E (encapsidation sequence), AUG (initiator codon of GAG), SD (splice donor) and DLS (dimer linkage structure) are underlined (see ref.8). Primer $tRNA^{Trp}$ (probe A) and the four synthetic DNA oligonucleotides (probes B to E) were 5' ^{32}P-labelled and used in hybridization assays in conditions reported in materials and methods. The probable involvement of different parts of the 5' 600 nt of RSV RNA in the dimer structure is indicated by ++ (extensive hybridization), + (positive hybridization) and − (no hybridization). Hybridization of DNA oligomer C was hardly detectable to both the monomer and the dimer forms of RSV RNA.
B) Incubation of RSV RNA of 680 nt (0,5 μg of heat denatured RNA per assay) was carried out in dimerization assay buffer (see materials and methods) with NCp12 and 10,000 cpm 5' ^{32}P-$tRNA^{Trp}$, or 50,000 cpm 5' ^{32}P DNA oligomer B or E, and 150,000 cpm of 5' ^{32}P DNA oligomer D.
Agarose gel electrophoresis was from top to bottom and autoradiography was for 6 hs. M and D indicate the monomer and dimer forms of RSV RNA.

Lanes 1−3.	Primer $tRNA^{Trp}$:	1/ control,	2/ 30 ng NCp12,	3/ 60 ng NCp12.
Lanes 4−6.	Oligomer B :	5/ control,	4/ 60 ng NCp12,	6/ 30 ng NCp12.
Lanes 7−9.	Oligomer D :	7/ control,	8/ 30 ng NCp12,	9/ 60 ng NCp12.
Lanes 10−12.	Oligomer E :	10/ control,	11/ 30 ng NCp12,	12/ 60 ng NCp12.

10 mM Tris-CL pH 7,0 for 5 min and subsequently analysed by agarose gel electrophoresis. As shown in figure 1 (lanes 7−10), thermal transition of RSV RNA dimer to monomer occured at 45−50°C, a temperature very dose to the Tm of 45−52°C measured for RSV 70S RNA (9).

Dimerization of RSV RNA catalysed by nucleocapsid protein NCp12

It has been noted in a previous report (21) that the viral protein, NCp12, catalysed the hybridization of replication primer $tRNA^{Trp}$ to the primer binding site (PBS) at positions 102−120 (8,21). Since this annealing reaction was always concomitant with the formation of dimeric RNA molecules, NC protein was suspected to activate RNA dimerization. Therefore dimerization of RSV monomeric RNA was monitored in absence and presence of NCp12 at 37°C and in a low ionic strength buffer.

In these ionic conditions spontaneous dimerization of RSV RNA monomer can occur but it is a slow process since less than 10% of RSV RNA was dimeric after 3 h at 37°C (figure 2, lane 2). In the presence of NCp12 the incubation time was limited to 15 mn and the RNA was deproteinized before electrophoresis in order to prevent any gel retardation effect due to bound NC protein molecules. Figure 2 (lanes 4 & 5) reports that NCp12 greatly activated RSV RNA dimerization at a degree depending

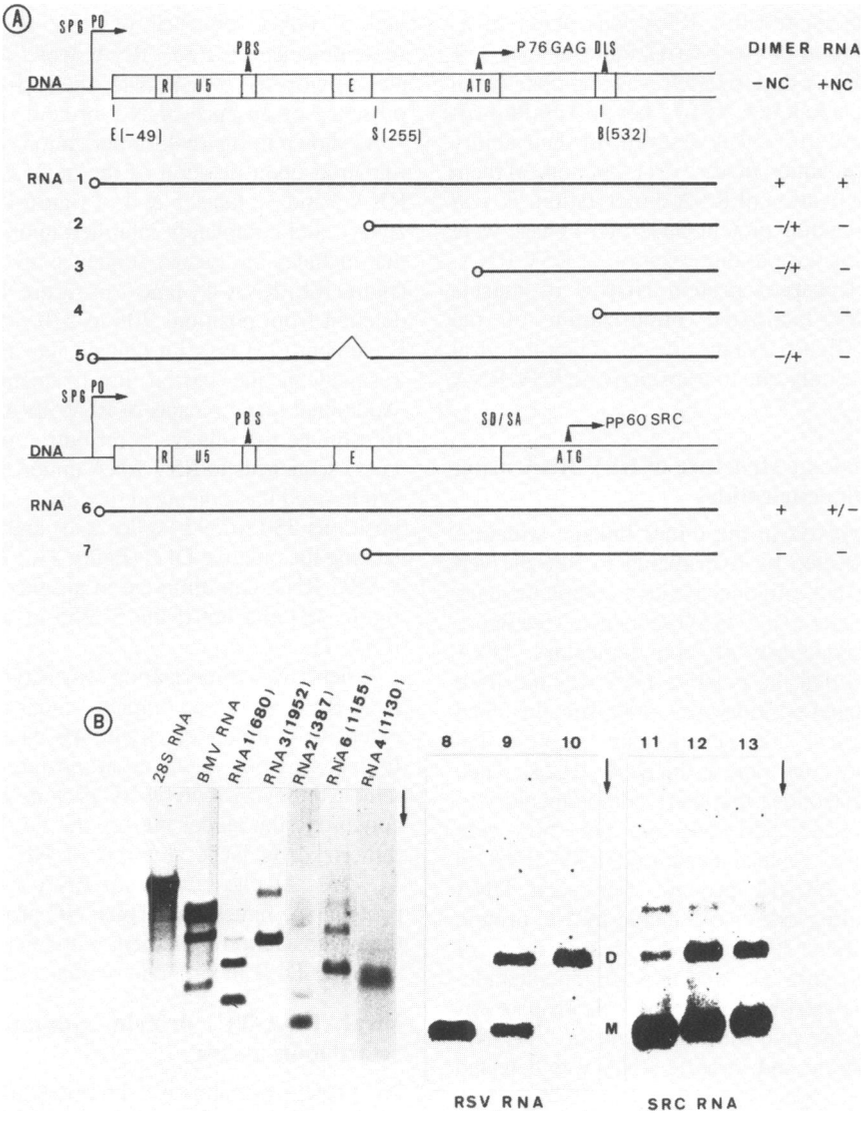

Figure 4: Mapping of the sequences required for RSV RNA dimerization by deletion mutagenesis.

A) RSV DNAs and SRC DNAs with different deletions were used for transcription with SP6 or T7 RNA polymerase (see methods). The 5′ 600 nt of interest are shematized : R (terminal repeat), U5 (5′ untranslated sequence), PBS (primer binding site), E (encapsidation sequence), ATG (initiator codon for $Pr76^{gag}$ synthesis), DLS (dimer linkage structure). For SRC DNA, SD/SA corresponds to the linkage of the v-SRC sequence to the 5′ 398 nt of RSV, ATG to the initiator codon for pp60src synthesis and the DLS is absent. The restriction sites of interest are indicated: EcoRI (−39), SacI (+255) and BamHI (+532).

Below the template DNAs are drawn the recombinant RNAs. The 5′ end is shown by an open circle and the deletion is indicated by a broken line. The size of each RNA species (1 to 7) used in the assays is indicated below :

RNA 1: pRSV LAD4 RNA of 680 nt with complete Leader.
RNA 2: pRSV LAD5 RNA of 387 nt with 5′ first 255 nt deleted.
RNA 3: pRSV LAD6 RNA of 1952 nt with 5′ first 372 nt deleted.
RNA 4: pRSV LAD9 RNA of 1130 nt with 5′ first 532 nt deleted.
RNA 5: pRSV LAD8 RNA of 620 nt with deletion from pos. 208−270.
RNA 6: pSRC1 RNA of 1155 nt with complete Leader.
RNA 7: pSRC2 RNA of 906 nt with 5′ first 255 nt deleted.

Dimerization assays were performed as described in figure 1 and 2 and with or without NC protein. Results are summarized on the right for each RNA. Presence of dimeric RNA is indicated by +, absence by −. Sign +/− indicates that no more than one third of the RNA was dimer.

B) Some of the RNA dimerization assays are shown. Agarose gel electrophoresis of RNA in native conditions and ethidium bromide staining were as described in methods.

Lanes 1 & 2 : Control marker 28 S rRNA and BMV RNAs.
Lanes 3 to 7 are spontaneous RNA dimerization (see methods).
Lane 3, RNA 1 with complete Leader.
Lane 4, RNA 3 with deletion of the 5′ first 372 nt.
Lane 5, RNA 2 with deletion of the 5′ first 255 nt.
Lane 6, RNA 6 with complete Leader but without DLS.
Lane 7, RNA 4 with deletion of the 5′ first 532 nt.
Lanes 8 to 13 are RNA dimerization activated by NCp12.
Lanes 8−10 : RSV RNA with complete Leader, 8/control, 9/ with 30 ng NCp12, 10/ with 80 ng NCp12.
Lanes 11−13: SRC RNA with complete Leader, 11/ control, 12/ with 50 ng NCp12, 13/ with 100 ng NCp12.

upon the amount of protein added : RNA was about 50% dimerized at an estimated ratio of one NCp12 molecule per 350 nt of RNA (lane 4), and RNA dimerization was complete when this ratio was increased to about one NCp12 per 140 nt (lane 5).

Other viral proteins were previously assayed for their ability to activate RSV RNA dimerization *in vitro* (21), but none of them showed any convincing activation of RNA dimerization. *E. coli* RecA and nucleic acid unwinding protein Gp32 of T4 phage were also incapable of promoting a rapid dimerization of RSV RNA. On the other hand nucleocapsid protein NCp10 of murine leukemia virus (MuLV) facilitated dimerization of the heterologous RSV RNA (figure 2, lanes 6−8). Then the viral NC protein is probably the only one to trans-activate RSV RNA dimerization.

Mapping of the dimer linkage structure of RSV RNA using complementary DNA oligonucleotides

RSV RNA sequences involved in the dimer linkage structure (DLS) were tentatively located by hybridizing to SP6-derived RNA a set of synthetic DNA oligonucleotides complementary to small domains of the Leader and 5′ GAG sequences (see figure 3A). To choose the sequences of complementary DNA oligonucleotides we noted that the putative DLS was localised by electron microscopy around positions 480−540 from the 5′end (7). In addition biochemical and genetic data suggest that sequences in the leader may contribute to the DLS (10,23). Then selected synthetic oligonucleotides that were complementary to positions 211−232, 364−383, 520−541 and 544−564 were 5′^{32}P labelled and annealed to heat denatured RSV RNA in absence or presence of NCp12 protein. For each DNA oligonucleotide two annealing assays conducted by NC protein were performed so that about 30% and 60%, respectively, of RSV RNA was dimeric (figure 3). ^{32}P-DNA oligonucleotide-RNA complexes were analysed by agarose gel electrophoresis in native conditions, and the hybridization level of ^{32}P-DNA oligonucleotide to monomeric and dimeric RNA was detected by autoradiography.

In agreement with our previous results (21) replication primer tRNA^{Trp} (probe A) hybridized to both the monomer and the dimer forms of RSV RNA (figure 3B, lanes 1−3) and this hybridization was well enhanced by NCp12 protein (lane 1 is control, lanes 2 & 3 are with NCp12). The four different DNA oligonucleotides B,C,D and E (figure 3A) were used in the hybridization assays at a probe to RNA molar ratio of 1 to 10 for probes B and E, and at an increased ratio of 1 to 3 for probes C and D because they poorly annealed to RSV RNA. Analyses indicated that probes B (lane 5, control ; lanes 4 & 6 with NC) and D (lane 7, control ; lanes 8 & 9 with NC) hybridized to both the monomeric (M) and dimeric (D) RNA while probe E (lane 10, control ; lanes 11 & 12 with NC) mostly hybridized to the monomeric RNA. Probe C poorly annealed to both forms of RSV RNA (not shown). Thus, and in agreement with data obtained by electron microscopy (7), these results supported the notion that sequences between positions 544−564 of RSV RNA most probably correspond to part or all of the DLS.

Cis-acting elements needed for RSV RNA dimerization in vitro

The next question to be addressed was : What are the sequences needed for the dimerization of RSV RNA ? In an attempt to identify such sequences deletions were made in the 380 nt leader and 5′ GAG sequence of RSV RNA. SP6 or T7 derived recombinants of RSV RNA were analysed by agarose gel electrophoresis to estimate their degree of dimerization in presence or absence of NC protein (figure 4 A & B).

As shown in figure 4, dimerization of RSV RNA was partially impaired upon deletion of the 5′ 255 nt or 372 nt (figure 4A, RNA 2 and 3 ; lanes 5 & 4 of figure 4B ; lane 3 is control RSV RNA), and completely inhibited upon removal of the 5′ 532 nt that includes the Leader sequence and part of the putative DLS (figure 4A, RNA 4 ; lane 7 of figure 4B). Subsequently a small deletion from positions 208 to 270 removing the encapsidation E element (23) was introduced into RSV RNA (RNA 5), and resulted in the partial impairement of RNA dimerization. Additional recombinants of RSV RNA were constructed in order to estimate the relative contribution of the leader and of the 5′ GAG sequences to RSV RNA dimerization. A v-SRC RNA was synthesized that contained the leader, the 5′GAG sequence from positions 380 to 397 (splice site) and the v-SRC sequence, but lacking the putative DLS (figure 4A ; RNA 6). This SP6-derived v-SRC RNA can dimerize in absence of NC protein (lane 6 of figure 4B) and not if the 5′ 255 nt were missing (figure 4A ; RNA 7).

When RNA dimerization was activated by NC protein, only RSV RNA with the complete leader and the putative DLS was extensively dimerized (figure 4B, lanes 8−10). Deletion of the E encapsidation element at positions 208−270 (figure 4A) inhibited dimerization by NC protein. Similarly the v-SRC RNA, containing the leader but not the DLS, was no more than 25% dimerized by large amounts of NC protein (figure 4B, lanes 11−13). All the other viral RNA were not dimerized by NC protein (see figure 4A). Thus, NC protein appeared to need both the encapsidation E element in the leader and the 5′end of GAG with the DLS in order to activate RSV RNA dimerization.

Viral RNA-NC protein complexes visualized by gel retardation assays

NC protein was shown to be bound to the genomic RNA dimer in virions (10−12) and to activate RNA dimerization (figure 2), and this prompted us to investigate the interactions between NC protein and RSV RNA by gel retardation assays. Incubation of RSV RNA with NC protein was carried out as before (see figure 4), and the NC-RNA complexes were analysed by agarose gel electrophoresis in a Tris-acetate buffer pH 7. RNA was visualized by ethidium bromide staining and NC protein by means of transferring the complexes onto nitrocellulose which were probed with anti-NCp12 serum (24,37).

As shown in figure 5, complexes contained both the RNA and the NC protein (see lanes 2−4 and 17−20 for examples). Upon addition of increasing amounts of NC to RSV RNA with the leader and 5′ GAG sequences, migration of the complexes was progressively retarded (lane 1, control; lanes 2−4, NC to RSV RNA ratio of 1 NC per 350 nt, per 175 nt and per 80 nt, respectively) and at a ratio of about 1 NC per 80 nt of RSV RNA, complexes poorly entered the agarose gel (lane 4). Similar gel retardation assays were performed with NC protein and BMV RNA, 28 S rRNA from CEF cells and SP6-derived v-Src RNA. Using SP6-derived v-SRC RNA, or RSV RNA lacking the 5′255 nt (see figure 4), retardation of the complexes still took place but it was less effective than with the complete RSV RNA made in vitro (lane 13, control ; lanes 14−16, with NC). Upon addition of NC protein to BMV RNA (lane 5, control ; lanes 6−8, with NC) and to the 28 S rRNA (lane 9, control ; lanes

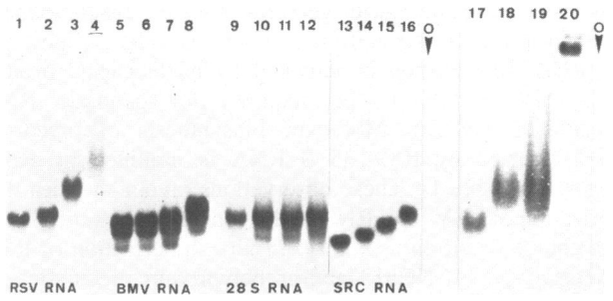

Figure 5: Visualization of NCp12-RSV RNA complexes by gel retardation assays. Complexes between RSV RNA of 2795 nt (pLAD4 cut at HindIII, see materials and methods)(0,5 μg), or BMV RNA (1μg), or 28 S rRNA (0,5 μg) or SRC RNA 6 of 2505 nt (pSRC1 cut at PvuI site, see materials and methods and figure 4A)(0,5 μg) and NCp12 were formed in conditions described for RNA dimerization (see methods). Agarose gel electrophoresis in a Tris-acetate buffer was as described in methods.

Complexes were visualized by ethidium bromide staining (lanes 1−16) or by immunodetection using anti-NCp12 serum (lanes 17−20).

The assays involved a control without NCp12 :

Lanes of control RNA: 1/ RSV RNA, 5/ BMV RNA, 9/ 28 S RNA, 13/ SRC RNA

NCp12 to RNA ratios of 1 NC per 350 nt, or 175 nt or 80 nt were used.

Lanes 2−4: RSV RNA. Lanes 6−8: BMV RNA. Lanes 10−12: 28 S rRNA. Lanes 14−16: SRC RNA.

Lanes 17−20. RSV RNA. Ratios of 1 NC per 700 nt, or 350 nt, or 175 nt, or 80 nt, respectively. These assays detected the NCp12 protein.

Note that the complexes stayed at the top of the gel when NCp12 to RSV RNA ratio was 20 (lanes 4 & 20) but not using RSV RNA lacking the 5'255nt (not shown) or SRC RNA.

10−12, with NC) almost no retardation of the complexes took place.

Additional retardation assays were performed with NC and RSV RNA lacking the leader and 18 S rRNA from CEF cells, and no retardation was observed (results not shown). Similarly other nucleic acid binding proteins like AMV reverse transcriptase and protein Gp32 from T4 phage were used and their binding to RSV RNA with the leader and 5'GAG sequences was monitored by this gel retardation assay. No retardation was observed although proteins were binding to RSV RNA in these conditions (results not shown). These results confirm that NC binds to RNA (10−15) which, in addition to activate dimerization of RSV RNA (figures 2 & 4), probably causes the formation of high molecular weight ribonucleo-protein complexes formed of NC molecules and RSV RNA with the leader and 5' GAG sequences. Such high molecular weight ribonucleo-protein complexes were not formed, or were labile, with RNA lacking either the leader or the 5' GAG (figure 5) although NC protein actually binds to these heterologous RNA (13,14).

Synthesis of polyprotein Pr76gag inhibited by dimerization of RSV RNA

In looking for a possible function of RSV RNA dimerization one might remember that, I) The 35S RNA monomer is both the mRNA for Pr76gag and Pr180gag-pol synthesis and the precursor to the 70S RNA dimer (1), II) both the 35 S RNA and the 70S RNA are found in RSV infected cells (1,9), and III) RSV RNA dimer interacts with NC protein molecules in the capsid (10,11).In addition a previous report showed that RSV RNA monomer or heat denatured RSV RNA dimer was efficiently translated in the Rabbit Reticulocyte lysate (RRL) while native RSV RNA dimer was not (10,24). So, we hypothesized that RSV RNA dimerization by NC protein could be a way to stop translation and direct encapsidation of the genome length RSV RNA into virion.

Table 1 : Dimerization of RSV RNA inhibits synthesis of precursor protein $Pr76^{gag}$

Messenger RNA	Control	Protein synthesis With NC to RNA ratio : 1 NC/350 nt	1 NC/175 nt
RSV 35 S RNA	100% (125 000 cpm)	21%	5%
RSV RNA 1 (complete 380 nt Leader)	100% (65 000 cpm)	26%	10%
RSV RNA 3 (without Leader, figure 4)	100% (31 000 cpm)	95%	80%
BMV RNA	100% (400 000 cpm)	95%	85%

RSV RNA and BMV RNA were first incubated with or without NC protein in conditions reported in methods and subsequently translated in the rabbit reticulocyte lysate at 20 μg/ml in conditions described in materials and methods. ^{35}S labelled proteins were analysed by SDS-PAGE electrophoresis and the radioactive bands corresponding to RSV $Pr76^{gag}$, and to BMV 100 and 30 Kda proteins were cut out and radioactivity counted in a scintillation counter. Values reported resulted from two separate experiments.

To test this hypothesis RSV RNA with the GAG coding sequence and with or without the 380 nt leader was incubated without NC or with NC at an NC to RNA ratio of 1 NC per 350 nt or per 175 nt. Then RNA was translated in the RRL system, protein analysed by polyacrylamide gel electrophoresis and amount of ^{35}S-labelled precursor $Pr76^{gag}$ was monitored by counting ^{35}S-Methionine incorporated into the protein. RSV 35 S RNA purified from virions and BMV RNA were used as positive and negative controls, respectively, of RNA translation in presence or absence of NC.

Table 1 reports that synthesis of ^{35}S-labelled $Pr76^{gag}$ was strongly inhibited by the NC protein when 35 S RNA or SP6-derived RSV RNA with the complete leader were used as messenger RNA. On the contrary protein synthesis is not or poorly inhibited by the NC protein when RSV RNA lacking the leader or BMV RNA were the messenger RNA.

DISCUSSION

The genetic material of RSV and probably of all retroviruses is an RNA dimer where two identical subunits are linked at their 5' end by a unique structure named Dimer Linkage Structure (DLS) (3,5−7). In addition $tRNA^{Trp}$, the replication primer for minus strand DNA synthesis, is hybridized to each RNA subunit at the primer binding site (PBS) at 102−120 nt from the 5' end (3,8).

Here we describe a system to study *in vitro* the dimer structure of RSV RNA as well as the dimerization process. SP6-derived RSV RNA had the capability to dimerize spontaneously (figure 1) indicating that a direct interaction between the two RNA molecules must take place without the involvement of any RNA or protein linker as previously suggested (39). In addition it is shown that nucleocapsid protein NCp12 catalyses dimerization of RSV RNA (figure 2).

Our first goal was to locate the DLS by hybridizing complementary DNA oligonucleotides to the monomer and dimer forms of RSV RNA. Assuming that the DLS corresponds to a stable RNA structure, DNA oligomers complementary to the DLS should anneal to the RNA monomer but not to the dimer

Table 2: Nucleotide sequence at and around the putative DLS of RSV RNA dimer

Positions are that of RSV PrC (8), and only mutations with respect to the sequence of PrC are indicated. All sequences shown have been published (see references 24,31) except that of ALV (Bieth and Darlix, unpublished results).

Positions	521	531	541	551	561
	.	.	.	.	.
PrC	CCGGGGUCCU	GGGAUCCCAU	UACCGCGG	CGCUAUCCCA	GCGG
PrB (HM)	----------	----------	---U----	----C-----	----
ALV	----------	----------	C--U----	----C-----	----
FSV	----------	----------	--------	-A--C-A-A-	----
MC29	--------A-	----------	--------	-A--CA----	----
UR2	----------	----------	C--U----	----C-----	----
Y73	----------	----------	--------	----C-----	----

form. Results of the DNA oligomer hybridization assays indicated that the DLS probably is located around position 550 from the 5'end of RSV RNA, which is in agreement with the data obtained by electron microscopy of the genomic 70 S RNA (7).

Are there some peculiar features of the RNA domain around position 550 which could explain why this probably corresponds to the DLS ? Schwartz and col. already noted the presence of a nearly perfect inverted repeat of 28 nt at positions 521−548 and which was named D.L.S. (8). This inverted repeat as well as nt 549−564 are well conserved in the avian retroviruses PrC, PrB(HM), FSV, MC29, UR2, Y73 (24,31) and in an avian leucosis virus (sequence not yet published, E. Bieth & J.L. Darlix) (see table 2). Also nt 521−564 correspond to a stable structure recovered from partial RNase digestion of RSV 35 S RNA (36). In addition a structure model for the 5′ 600 nt of RSV RNA has been elaborated by means of computer analysis of sequencing data of RSV RNA secondary structures (24). In this model the putative DLS is within an hairpin structure with bulged loops located at positions 500 to 618 and rich in G and C residues (24). How two of this hairpin structures can interact in the dimer linkage structure is still a matter of speculation. Nevertheless it must be pointed out that the putative DLS which probably involves 50 nt for each RNA molecule (7) is denatured at 55°C (9). Hence, the presence of unusual RNA conformations like pseudoknots or non canonical base pairing interactions can be suspected. This issue requires an extensive study of RSV RNA interstrand interactions at the level of the DLS, and this is presently in progress.

The results of RSV RNA dimerization obtained from deletion mutageneses in the Leader and in 5′ GAG sequence suggest that in addition to the DLS a sequence in the Leader probably is important for an extensive RNA dimerization (figures 3&4). This sequence has been mapped in the E domain known to be required for RSV genomic RNA encapsidation (23). Thus two domains apppear to be necessary for RSV RNA dimerization that are the 5′ GAG with the DLS and the E element. However we do not know if these two domains are sufficient for RSV RNA dimerization and clearly more work is needed to precisely characterize the sequences necessary and sufficient for RSV RNA dimerization.

If RNA dimerization is a rapid process in RSV infected cells, then the observation that it is solely activated by nucleocapsid protein NCp12 appears of importance (figure 2) (21). NCp12 could function by binding to specific RNA structures as indicated by its binding sites at and around the DLS of the genomic dimer RNA in virions (10,11). Once NCp12 molecules are bound to a limited number of sites on the RSV RNA monomer, they may unwind the RNA structures thereby facilitating interstrand base pairing interactions and the formation of genomic RNA dimer. This is presently under investigation *in vitro*.

What might be the functional significance of RSV RNA dimerization ? RSV RNA dimerization is dependent upon cis-acting sequences required for genomic RNA encapsidation, the E domain in the Leader and the 5′ GAG sequences (figure 3). RSV RNA dimerization is activated by nucleocapsid protein NCp12 also known to be required for genomic RNA encapsidation (18,20). Moreover biosynthesis of precursor $Pr76^{gag}$ directed by RSV 35 S RNA is inhibited by RNA dimerization (table 1). These observations favour the idea that dimerization of RSV 35 S RNA and its encapsidation are linked in the course of virion assembly. *In vivo* dimerization of RSV RNA catalysed by NCp12 and/or polyprotein precursor gag should stop translation of the 35 S RNA and then should render it available for encapsidation. Unique RNA structures possibly present at and around the DLS of RSV RNA dimer would then be recognized as the packaging signal(s). Last but not least, the association between a large number of NC protein molecules and RSV dimeric RNA may well form a circular chromatin like structure in the capsid (40) which then could facilitate the synthesis of a complete proviral DNA molecule by reverse transcriptase (41).

ACKNOWLEDGEMENTS

We wish to thank R. Katz and M. Stoltzfus for providing the pDL12 and the pSRC plasmids. In addition, we thank A.C. Prats, M. Gazeau, P.F. Spahr, R. Devoret and H.V. Kolbe (Transgene) for purified NCp10 protein of MuLV, Reverse Transcriptase of AMV, CAp27 and MAp19 of RSV, *E.coli* RecA and CAp24 and MAp17 of HIV-1, respectively. This work was supported by a grant to J.L. Darlix from the French programme against AIDS (ANRS).

Abbreviations used: RSV, Rous Sarcoma virus ; BMV, Brome mosaic virus; nt, nucleotide ; DLS, dimer linkage structure ; PBS, primer binding site ; NC, nucleocapsid protein; CA, capsid protein.

REFERENCES

1. Leong, J.A., Garapin, A.C., Jackson, N., Fanshier, L. Levinson, W. & Bishop, J.M. (1972) J. Virol. *9*, 891−902.
2. Fan, H. and D. Baltimore. (1973) J. Mol. Biol. *80*, 93−117.
3. Coffin, J.M. (1984) The structure of the retroviral genome. p. 261−368, In R. Weiss, N. Teich, H. Varmus and G. Coffin (ed.), RNA Tumor viruses, vol. 1. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.(USA).
4. Pugatsch, T. & Stacey, D. (1983) Virology *128*, 505−511.
5. Bender, W. and Davidson, N. (1976) Cell, *7*, 595−607.
6. Bender, W., Chien, Y.H., Chattopadhyay, S., Vogt, P.K., Gardner, M.B. and Davidson, N. (1978) J. Virol., *25*, 888−896.
7. Murti, K.G., Bondurant, M. and Tereba, A. (1981) J. Virol., *37*, 411−419.
8. Schwartz, D., Tizard, R. & Gilbert, W. (1983) Cell *32*, 853−869.
9. Stoltzfus, M. & Snyder, P. (1975) J. Virol. *16*, 1161−1170.
10. Darlix, J.L. & Spahr, P.F. (1982) J. Mol. Biol. *160*, 147−161.
11. Meric, C., Darlix, J.L. & Spahr, P.F. (1984) J. Mol. Biol.*173*, 531−538.
12. Darlix, J.L., Meric,C. & Spahr, P.F. (1985) Viral messenger RNA p355−372, Y. Becker (ed), Boston (USA).
13. Bailey, B.J. & Smith, J.M. (1979) Nucleic Acid Res. *7*, 2055−2072.
14. Sykora, K.W. & Moelling, K. (1981) J. Gen. Virol. *55*, 379−391.
15. Leis, J. & Jentoft, J. (1983) J. Virol. *48*, 361−369.
16. Covey, S.N. (1986) Nucleic Acid Res. *14*, 623−633.
17. South, T.L., Kim, B. & Summers, M.F. (1989) J. Am. Chem. Soc. *111*, 395−396.
18. Meric, C. & Spahr, P.F. (1986) J. Virol. *60*, 450−459.
19. Gorelick, R. J., L. E. Henderson, J. P. Hanser and A. Rein (1988) Proc. Natl. Acad. Sci. USA, *85*, 8420−8424.
20. Meric, C. and S. Goff. (1989) J. Virol., *63*, 1558−1568.
21. Prats, A.C., L. Sarih, C. Gabus, S. Litvak, G. Keith and J.L. Darlix. (1988) EMBO J. *7*, 1777−1783.
22. Panganiban, A. & Fiore, D. (1988) Science *241*, 1064−1069.

23. Katz, R., Terry, R.W. & Skalka, A.M. (1986) J. Virol. *59*, 163–167.
24. Darlix, J.L. (1986) J. Mol. Biol. *189*, 421–434.
25. Chang, L.J. & Stoltzfus, M. (1985) J. Virol. *53*, 969–972.
26. Maniatis, T., Fritsch, E.F., Sambrook, J. (1982) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.(USA).
27. Moelling, K., Sykora, K.W., Dittmar, K.E., Scott, A. and Watson, K.F. (1979) J. Biol. Chem., *254*, 3738–3742.
28. Fournier, M., Dorizzi, M., Serjer, C. & Labouesse, G. (1976) Biochimie *58*, 1159–1165.
29. Darlix, J.L., Spahr, P.F., Bromley, P.A. and Jaton, J.C. (1979) J. Virol., *29*, 597–611.
30. Zuker, M. and Stiegler, P. (1981) Nucl. Acids Res., *9*, 133–148.
31. Van Beveren, C., J. Coffin and S. Hughes. (1985) Appendixes p. 559–764. In R. Weiss, N. Teich, H. Varmus & J. Coffin (ed), RNA Tumor viruses, vol. 2 Cold Spring Harbor Laboratory, Cold Spring Harbor. N.Y. USA.
32. Munroe, S.H. & Duthie, R.S. (1986) Nucleic Acid Res. *14*, 8447–8465.
33. Ganoza, M.C., Kafoid, E.C., Marliere, P. & Lowis, B.G. (1987) Nucleic Acid Res. *15*, 345–359.
34. Darlix, J.L., Zuker, M. & Spahr, P.F. (1982) Nucleic Acid Res. *10*, 5183–5196.
35. Darlix, J.L., P.F. Spahr and P.A. Bromley. (1978) Analysis of Rous Sarcoma virus (RSV) RNA structure by means of specific nucleases. Virology, *90*, 317–329.
36. Darlix, J.L., Schwager, M., Spahr, P.F. & Bromley, P. (1980) Nucleic Acid Res. *8*, 3335–3354.
37. Prats, A.C., Roy, C., Wang, P., Erard, M., Housset, V., Gabus, C., Paoletti, C. & Darlix, J.L. J. Virol (in press).
38. Burnette, W.G. (1981) Anal Biochem. *112*, 195–203.
39. Kung, H.J., Hu, S., Bender, W., Bailey, J.M., Davidson, N., Nicolson, M.O. and R. M. Mc Allister. (1976) Cell, *7*, 609–620.
40. Chen, M.J., Garon, C. & Papas, T.S. (1980) Proc. Nat. Acad. Sci. (USA) *77*, 1296–1300.
41. Boone, L.R. & Skalka, A.M. (1981) J. Virol. *37*, 109–116.