Nucleic Acids Research, Vol. 19, No. 13 3533–3541

Viral RNA annealing activities of the nucleocapsid protein of Moloney murine leukemia virus are zinc independent

Anne-Catherine Prats, Valérie Housset, Gérard de Billy, Fabrice Cornille[1], Hervé Prats, Bernard Roques[1] and Jean-Luc Darlix*

Centre de Recherche de Biochimie et Génétique Cellulaires du CNRS, 118 route de Narbonne, 31062 Toulouse cedex and [1]Université Paris V, 4, avenue de l'Observatoire, 75006 Paris, France

Received April 18, 1991; Revised and Accepted June 12, 1991

ABSTRACT

The zinc fingers of retroviral *gag* nucleocapsid proteins (NC) are required for the specific packaging of the dimeric RNA genome into virions. *In vitro,* NC proteins activate both dimerization of viral RNA and annealing of the replication primer tRNA onto viral RNA, two reactions necessary for the production of infectious virions. In this study the role of the zinc finger of Moloney murine leukemia virus (MoMuLV) NCp10 in RNA binding and annealing activities was investigated through modification or replacement of residues involved in zinc coordination. These alterations did not affect the ability of NCp10 to bind RNA and promote RNA annealing *in vitro*, despite a complete loss of zinc affinity. However mutation of two conserved lysine residues adjacent to the finger motif reduced both RNA binding and annealing activities of NCp10. These findings suggest that the complexed NC zinc finger is not directly involved in RNA-protein interactions but more probably in a zinc dependent conformation of NC protein modulating viral protein-protein interactions, essential to the process of viral RNA selection and virion assembly. Then the NC zinc finger may co-operate to select the viral RNA genome to be packaged into virions.

INTRODUCTION

In the virion core of all known retroviruses, the genome is present as a 70S complex containing two identical unspliced viral RNA molecules. Two major non covalent RNA-RNA interactions take place close to the 5′ end of the viral RNA. One involves the joining of the two viral RNA molecules and has been named dimer linkage structure (DLS) (1). The other corresponds to the annealing of the replication primer tRNA of cellular origin at the primer binding site of the viral RNA (2,3). The RNA genome is also tightly associated with nucleocapsid (NC) protein molecules, a maturation product of the *gag* polyprotein (4,5). The highly basic NC protein is conserved among retroviruses, and contains one or two motifs of the form Cys-X_2-Cys-X_4-His-X_4-Cys (CCHC). Recently we and others have shown that the unique CCHC box of the MoMuLV NCp10 protein coordinates one zinc ion (6,7,8). Therefore this motif can be referred to as a zinc finger (9).

Point mutations that destroy the zinc finger of MoMuLV NC protein, or change the aromatic residues within the finger motif, result in a strong impairment of genomic RNA packaging (10,11). At the same time some of these MoMuLV NC mutants encapsidate various cellular RNAs indicating that the mutated NCs have lost their ability to select and encapsidate MoMuLV genomic RNA. Similarly mutations in the first zinc finger of Rous sarcoma virus (RSV) and human immunodeficiency virus (HIV) NCs impair the packaging of viral genomic RNA (12,13,14). These data indicate that the NC zinc finger probably controls the encapsidation of viral genomic RNA. Also, packaging of the genomic RNA necessitates a *cis*-element which has been identified and located close to the 5′ end of the viral genome in avian, mammalian and human retroviruses (15,16,17,18). In MoMuLV, this element named *Psi* has been mapped between nucleotides 215 and 565 of the genomic RNA (16). Presence of *Psi* within an RNA, but only in the sense orientation, is necessary and sufficient to promote RNA packaging (19). This property of *Psi*, which has led to the design and extensive uses of retroviral vectors, suggests that the packaging signal probably involves a cis-acting RNA structure.

In vitro, the mature NC protein shows preferential binding to single stranded nucleic acids, and the highest affinity for the viral dimeric 70S RNA (20,21,22). NC protein from RSV, MoMuLV, and HIV also exhibits a nucleic acid annealing activity which facilitates dimerization of retroviral *Psi*+ RNA and annealing of the replication primer tRNA onto the viral RNA (23,24,25,26). While the tRNA annealing reaction is probably required only to initiate reverse-transcription of the genome, viral RNA dimerization could be involved in two critical processes, RNA packaging and reverse transcription. Indeed, retroviral RNA dimerization requires *cis* elements located within the encapsidation element (25,26,27). Furthermore mutant viruses partially defective in packaging contain low amounts of genomic RNA which is mostly dimeric (10,11), and dimeric RNA has

* To whom correspondence should be addressed at LaboRetro (INSERM), Ecole Normale Supérieure de Lyon, 46, allée d'Italie, 69364 Lyon cedex 07, France

also been detected in infected cells (28). These observations suggest that retroviral RNA is probably packaged in the dimer form. In addition the presence of the DLS may be important for the template switching by reverse transcriptase during the proviral cDNA synthesis (29).

In this work we have investigated the role of the zinc finger of MoMuLV NCp10 in viral RNA binding, dimerization and primer tRNA annealing onto MoMuLV RNA. NCp10 was either mutated within and next to the finger motif and purified from *E.coli*, or alkylated on the cysteines of the finger using a chemically synthesized protein. Mutant and modified NCp10 were assayed for their ability to bind zinc and viral RNA, as well as to activate viral RNA dimerization and primer $tRNA^{Pro}$ annealing *in vitro*. Data clearly show that MoMuLV RNA binding and annealing activities of NCp10 protein do not require the coordination of a zinc ion.

MATERIALS AND METHODS

Plasmid constructions and NCp10 mutagenesis

pRFC10: A DNA fragment containing the coding sequence of human basic fibroblast growth factor (bFGF) between *NdeI* and *Hind III* sites was generated by PCR (H. Prats, manuscript in preparation) and introduced into the expression vector pET3 (30). The recombinant plasmid was digested by enzymes *BglII* and *EcoRI*, and the generated fragment was introduced into the high copy number vector pSP65 opened by enzymes *EcoRI* and *BamHI*. The resulting plasmid, pRFC10, allows expression of protein coding sequences introduced into the *NdeI* site.

pACP10W: NCp10 DNA was amplified by polymerase chain reaction (PCR) in order to obtain a fragment with the NCp10 open reading frame between two *NdeI* sites. The template was plasmid pMLVAC-108 containing the MoMuLV *gag* gene (27,31), and primers had the sequences AAACATATGCTAC AGGAGGGAGGTCTG and AAACATATGGCCACTGTCG TTAGT. They are respectively complementary to the 3′ end and homologous to the 5′ end of the NCp10 open reading frame (initiation and stop codons are in bold characters, *NdeI* site is underlined). 10 ng of *BamHI*-linearized pMLVAC-108 DNA was amplified by 25 cycles of PCR (1 min at 95°C, 2 min at 37°C, 2 min at 72°C) in presence of both primers (10 μM each) and of *Taq* DNA polymerase (4 units). The 180 nt-length DNA fragment obtained was purified from low melting agarose, digested by enzyme *NdeI* and cloned into the *NdeI* site of plasmid pRFC10.

pACP10C1, C2, C3, T and L: plasmid pACP10W was digested by enzyme *HindIII* and the two fragments obtained were purified from low melting agarose gel. The vector fragment was dephosphorylated. The fragment containing NCp10 sequence was ligated overnight in diluted conditions (5μg/ml) in order to be circularized. 10 ng of this template was amplified by PCR in the same conditions as for pACP10W but using 'back to back' primers (see fig.1). Amplified DNA was treated by Klenow enzyme and T4 kinase to restore blunt ends, circularized by ligation as above, digested by enzyme *HindIII*, and inserted into pACP10W vector *HindIII* fragment (fig.1B). Mutations were screened by DNA sequencing using 'sequenase' (a modified T7 DNA polymerase provided by USB). For C1, C2 and T, DNA primers were GCGATCGAGTTGGGACC (antisense) and G-ACCAGTCTGCCTAC$_A$TGCAAAGAA (for C1 and C2, degenerate primer with A or T at position 16; mutations are in bold characters) and GACCAGTGTGCCTCCTGCAAA-GAAAAGGGGCACGGGGCT (for T). For C3 and L, DNA primers were TTTAGCCCAGTGCCCCTTTTC (antisense), G-ATGGTCCCAAGAAACCACGAGGA (for C3), and G-ATTGTCCCACGATACCACGAGGA (for L). For the construction of pACP10C3, the PCR template was pACP10C2 in order to obtain the three cysteine residues mutated together.

pMLVAC20: a *SacI-PstI* DNA fragment containing MoMuLV genome from nt −30 to nt 567 was inserted into Blue Scribe vector.

NCp10 production in *E.coli*

NCp10 was produced using the T7 expression system reported by Studier *et al.* (30). The plasmids pACP10 described above were introduced into the *E.coli* strain BL21(pLysS) (30). This strain contains the gene coding for the T7 RNA polymerase under control of promoter *lac*UV5. Moreover it contains the plasmid pLysS bearing the gene of resistance to chloramphenicol and the gene coding for T7 lysozyme. BL21(pLysS) containing plasmids pACP10 was grown at 30°C in LB medium (32) in presence of ampicillin and chloramphenicol. When cultures reached 0.3 to 0.6 OD_{600}, NCp10 production was induced by adding 1mM isopropyl-thio-β-galactoside (IPTG). After 2 to 5h of induction at 30°C, bacteria were chilled in ice and centrifuged 5 min at 5000 rpm. For analytical assays, pellets were resuspended directly in sample buffer (60mM Tris pH6.8, 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.2% bromophenol blue), in 1/10 of culture volume, then sonicated. Proteins were analyzed by electrophoresis on Tricine-SDS-10% polyacrylamide gel (PA-GE; 33) and Coomassie staining. Tricine, used as the trailing ion, allows resolution of small proteins at lower acrylamide concentrations than in glycine-SDS-PAGE systems. For preparative assays, pellets from 0.5 to 1l of bacteria cultures were resuspended in 5 ml of degased buffer containing 50mM Tris pH8.5, 0.1M NaCl, 5% glycerol, 5mM dithiothreitol (DTT), 1mM $MnCl_2$. Bacteria suspensions were conserved at −20°C.

NCp10 purification

0.5 to 1l of bacteria BL21(pLysS) containing plasmids pACP10 were grown, induced, centrifuged and resuspended as described above. Bacteria were lysed by a freezing-thawing step followed by incubation at 30°C for 30 min, in presence of 1mM phenyl-methyl-sulfonyl-fluorid. DNA was broken by incubation for 15 min at 37°C with DNase I at 50 μg/ml. The suspension was diluted in 10 ml of degassed Tris 10mM pH8.5, 5% glycerol, 1mM DTT and centrifuged at 100000g (35 krpm in the Ti60 rotor). The supernatant (S100) was collected.

NCp10 purification was performed in two steps using an FPLC apparatus (Pharmacia). All buffers were degassed under argon to avoid NCp10 oxidation. For the first step of purification, the S100 was loaded onto a 1ml heparine sepharose column, at a flow of 0.5 to 1ml/min, in 10mM Tris pH8.5, 0.1M NaCl, 5% glycerol, 1mM DTT. The column was washed in the same buffer until OD_{280} reached the base line. Proteins were eluted at 1ml/min by a 20 ml gradient from 0.1M to 2M NaCl. Fractions of 0.5 or 1ml were collected, and 5−10 μl of each fraction was analyzed by electrophoresis as described below. For the second step of purification, fractions containing NCp10 were diluted 3−4 times in 10mM Tris pH8.5, 5% glycerol, 1mM dithiothreitol so that the final NaCl concentration was 0.1M. These fractions

were loaded onto cation exchanger MonoS or S-sepharose column (Pharmacia). Proteins were eluted in the same conditions as for HS column. MonoS column has the advantage of concentrating the proteins, when compared with S-sepharose. When proteins were not pure enough, one or both steps could be repeated.

NCp10 was also purified from MoMuLV virions: virions were produced and processed as described before (27), and the viral lysate was directly loaded onto MonoS or S-Sepharose column as above. CAp30 (negative control) was purified from the viral lysate by phosphocellulose chromatography (34).

NCp10 peptide synthesis and thiol alkylation

NCp10 peptide was assembled on phenylacetamidomethyl resin on a semi-automated multichannel peptide synthesizer NEOSYSTEM NPS 4000 as previously described (8). Thiol groups were alkylated with iodoacetamide as previously described (35). Alkylated NCp10 was purified from reagents by reverse phase HPLC. No free thiol could be detected by titration by 5,5'-dithiobio(2-nitrobenzoic acid).

Zinc binding assay

The zinc binding assay was derived from that of Schiff et al. (36). For western blots, proteins were separated by tricine-SDS-PAGE, then transferred onto 0.1μ nitrocellulose membrane at 75 mA for 1h in 30% methanol, 3g/l Tris, 17g/l glycine. These conditions prevented NCp10 from going throughout the membrane. The filter was incubated for 30 min at 20° in 50mM Tris pH7, 50mM NaCl, 1mM DTT for protein

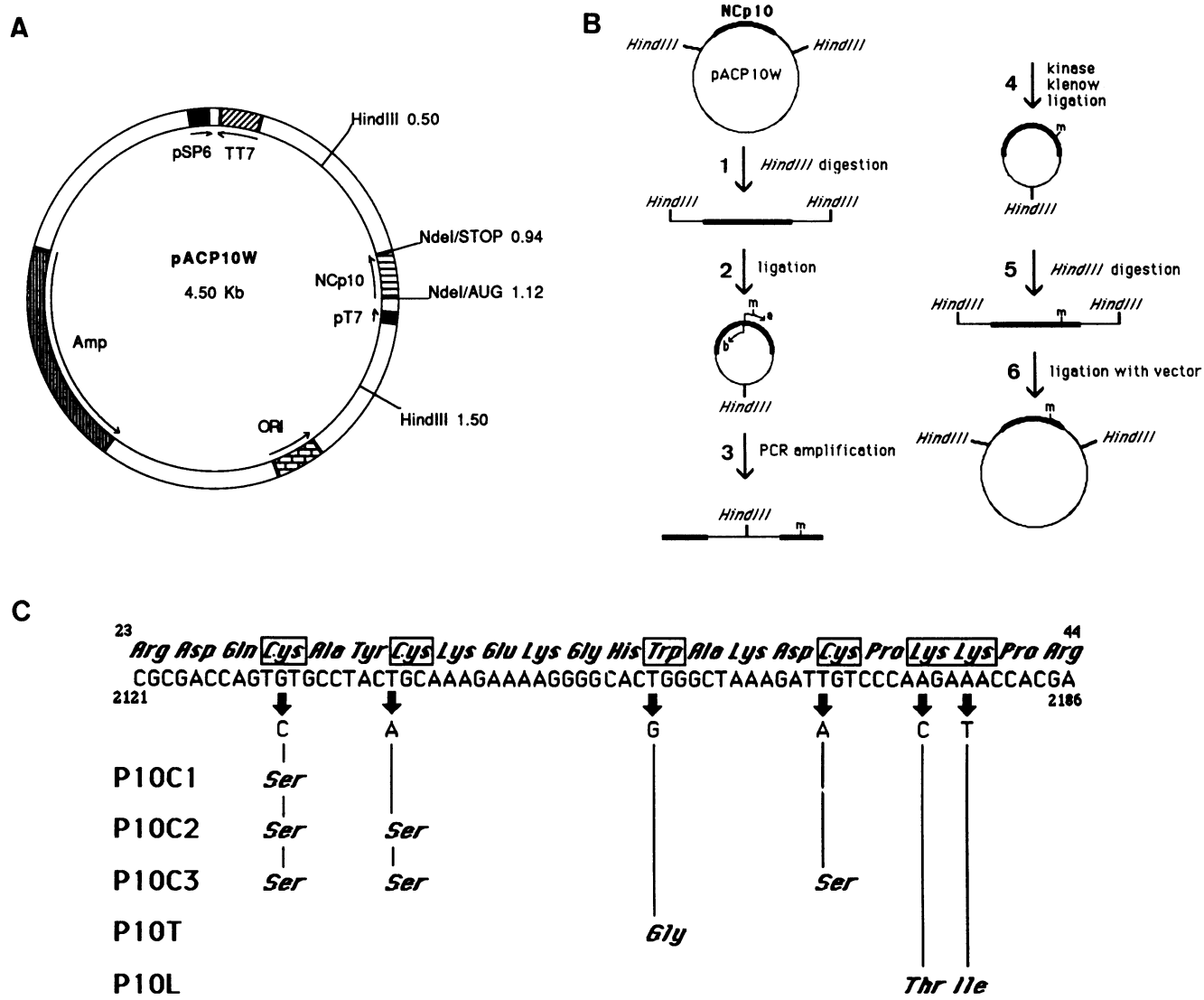

Figure 1. Cloning and mutagenesis of MoMuLV NCp10 coding sequence in a T7 expression vector. A) A DNA fragment containing the coding sequence of NCp10 was generated by PCR and introduced into the *NdeI* site of T7 expression vector pRFC10, as described in Materials and Methods. In the resulting plasmid pACP10W, NCp10 DNA is under control of T7 promoter, allowing expression of NCp10 in a bacteria strain producing T7 RNA polymerase (30). The restrictions sites used for pACP10W construction and NCp10 DNA mutagenesis, the gene of resistance to ampicillin (*amp*), the origin of plasmid replication (*ori*), T7 promoter (pT7) and terminator (TT7) are indicated. B) The procedure used to generate point mutations in NCp10 zinc finger is schematized. 1) Plasmid pACP10W was digested by *HindIII*, 2) The generated 0.8 kb-lenth fragment was circularized by ligation. 3) The circular template was amplified using 'back to back' primers (a and b) by 25 cycles of PCR , as described in Materials and Methods. One of the primers bears the mutation (m). 4) The inverted fragment obtained was treated by T4 kinase and DNA polymerase I (Klenow), then ligated as above. 5) The ligation product was digested by *HindIII* . 6) The mutated 0.8 kb-length fragment was ligated with the *HindIII*-digested vector fragment from plasmid pACP10W. C) Nucleotide and aminoacid sequences of the region of NCp10 containing the CCHC box are represented. The aminoacids residues mutated in this report are in bold characters. Mutations are indicated, as well as the five plasmids pACP10C1, C2, C3, T and L generated by the procedure described in figure 1B.

renaturation, probed with $^{65}ZnCl_2$ at 2μCi/ml (15μM) for 30 min at 37°C and washed for 30 min at 20°C in the same buffer. Then the filter was dried and autoradiographed. For dot blots, proteins were spotted directly onto nitrocellulose membrane which was probed like western blots.

RNA binding assay

The RNA binding assay was derived from the 'southwestern' assay described by Bowen *et al.* (37), except that the proteins were directly spotted onto 0.1μ nitrocellulose membrane. 20μl of increasing dilutions of each NCp10 species (diluted in 10mM Tris pH 8.5, 0.3M NaCl, 5% glycerol, 5mM DTT) was dot blotted. The membrane was blocked for 30 min at 20°C in 20mM Tris HCl pH7, 50mM NaCl, 2mM $MgCl_2$, 1mM DTT, 20μg/ml bovine serumalbumine, then incubated with ^{32}P labelled MuLV RNA (5×10^5 cpm/ml) for 30 min at 20°C in the same buffer plus 10μg/ml *E.coli* tRNA, and finally washed for 15 min in the same conditions. After autoradiography, each spot was cut and the bound radioactivity counted.

Assay for RNA dimerization and tRNAPro hybridization

MoMuLV RNA was synthesized *in vitro* by T7 RNA polymerase and purified as described before (27) from pMLVAC-20 DNA template linearized with enzyme *HindIII* , resulting in a 0.6 kb-length RNA containing the dimerization sequence (27). $tRNA^{Pro}$ purified from beef liver was a gift from G. Keith. It was 5'-labelled (2×10^6 cpm/μg) by T4 kinase in presence of $\gamma^{32}P$ ATP, then repurified from 6% polyacrylamide-urea gel (23).

The assay was performed as previously reported (27) with a few modifications. RNA (0.3 mg) was denatured in H_2O at 100°C for 2 min and incubated at 37°C for 5 min in 10μl of Tris 20mM pH6.5, 60mM NaCl, 5mM DTT with ^{32}P-$tRNA^{Pro}$ (3×10^4 cpm) and 12 to 200 ng of purified NCp10. After the reaction, NCp10 was extracted by 10μl of phenol equilibrated with 50mM Tris pH 7.5, 0.3M NaCl, 0.1% SDS, and the aquous phase was analyzed by electrophoresis on 0.8% agarose gel in native conditions. Gel and migration buffer (Tris borate) contained ethidium bromide at 0.5 μg/ml, allowing us to visualize the RNA. Then the gel was fixed for 5 min in 5% TCA, dried and autoradiographed to estimate $tRNA^{Pro}$ annealing to the viral RNA containing the PBS.

RESULTS

Expression in E.coli and purification of MoMuLV NCp10 mutant proteins

A DNA fragment containing the exact open reading frame coding for mature NCp10 flanked by two *NdeI* sites was generated by

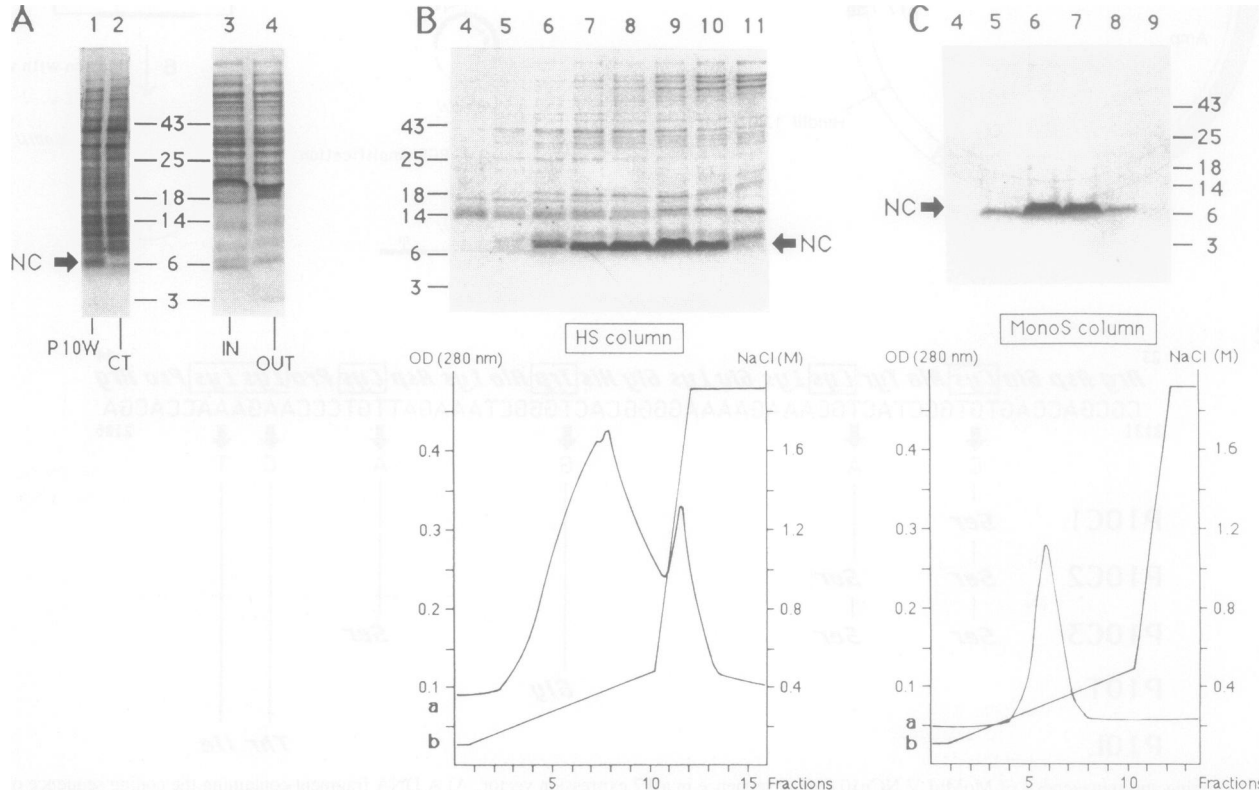

Figure 2. Production and purification of MoMuLV NCp10. Bacteria BL21(pLysS) (30).containing the plasmids pACP10 were induced by IPTG, lysed and centrifuged at 100000g (Materials and Methods). The S100 was loaded onto heparin-sepharose (HS) column and NCp10 eluted by NaCl gradient. NCp10-containing fractions were loaded onto MonoS column (cation exchanger), and NCp10 eluted again by NaCl gradient. Proteins were analyzed by Tricine-SDS-10% PAGE (33), a procedure allowing better migration of NCp10 (6.3 kda). A) Lane 1: lysate from bacteria BL21(pLysS) containing plasmid pACP10W. Lane 2: control lysate from bacteria containing the vector pET3 (30). Lane 3: S100. Lane 4: output from HS column. B) Elution of NCp10 from HS was performed by a 0.1M to 2M NaCl gradient. Line a is OD_{280}, line b is NaCl concentration. Analysis of HS fractions 3 to11 by PAGE is shown above HS diagram. Arrow indicates migration of NCp10. C) HS fractions 6 to 9 were pooled, diluted to 0.1M NaCl and loaded onto cation exchanger MonoS column. Elution of NCp10 from MonoS was performed as for HS. Analysis of MonoS fractions 5 to 10 is shown above the diagram.

polymerase chain reaction (PCR) and introduced into the T7 expression vector pRFC10 (fig. 1A).

To introduce mutations into the zinc finger of NCp10, we used a procedure (described in detail in figure 1B) based on PCR amplification of circular DNA templates (38). This procedure allowed us to mutate the NCp10 DNA sequence in and around the finger. As shown in figure 1C, one, two or three cysteine residues (Cys26, 29 and 39) of the finger were changed to serine residues. The tryptophane residue (Trp35) was changed to glycine and the two conserved lysine residues (Lys41 and 42) located downstream from the zinc finger were changed to threonine and isoleucine residues respectively.

The different DNA constructs coding for wild type or mutant NCp10 were introduced into BL21(pLysS), an *E.coli* strain expressing T7 RNA polymerase inducible by IPTG (30). After induction, NCp10 production was analyzed by PAGE, followed by coomassie blue staining or Western immunodetection (not shown). Large amounts of NCp10 were observed in induced cells (fig. 2A, lane 1). After cell lysis and membrane elimination most of NCp10 was found in the supernatant S100 (fig. 2A, lane 3). The S100 was chromatographed through an heparin sepharose (HS) column and the fractions containing NCp10 (eluted at 0.3M NaCl, fig. 2B) were re-chromatographed through a cation exchanger MonoS column. NCp10 was again eluted at 0.3M NaCl, giving a final concentration reaching 0.5 to 1 mg/ml (fig. 2C). This procedure allowed a rapid and efficient recovery of wild type and mutant NCp10 proteins in large amounts (several mg of NC protein from 1 liter of bacteria culture).

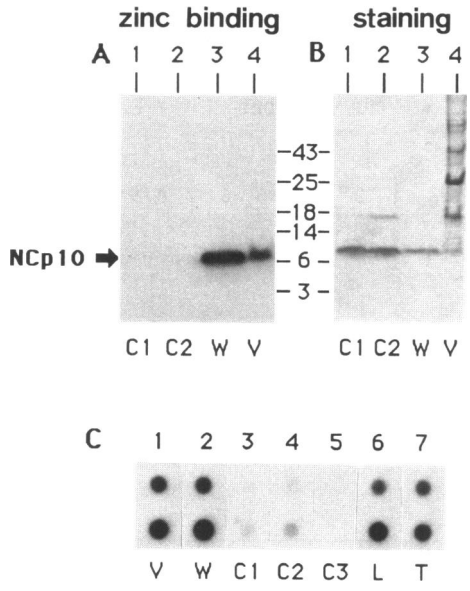

Figure 3. Binding of $^{65}ZnCl_2$ to wild type and mutant NCp10. A) **Zinc binding assay** on western blot. Proteins were electrotransferred from tricine-SDS-polyacrylamide gel onto nitrocellulose membrane, then probed with $^{65}ZnCl_2$ as decribed in Materials and Methods and autoradiographed for 24h. Lane 1: 1 μg NCp10-C1; lane 2: 1 μg NCp10-C2; lane 3: 0.5 μg wild type NCp10; lane 4: control with MoMuLV proteins. The presence of the proteins on the membrane was controlled by ponceau red staining. B) **Coomassie staining** of a polyacrylamide gel with the same samples as in A). C) **Zinc binding assay** on dot blot. Proteins were directly spotted onto nitrocellulose membrane and probed with $^{65}ZnCl_2$ as in A). 0.4 and 1 μg of each NCp10 species was spotted (upper and lower spots, respectively). 1/ Wild type NCp10 from virions. 2/ Wild type NCp10 from *E.coli*. 3/ NCp10-C1. 4/ NCp10-C2. 5/ NCp10-C3. 6/ NCp10-L. 7/ NCp10-T. Size standards are indicated, as well as the position of NCp10 migration. Under each lane or spot is indicated the NCp10 species used.

Binding of $^{65}ZnCl_2$ to wild type and mutant NC proteins

The finger motif of synthetic NCp10 has been shown to coordinate a cobalt or zinc ion with a high affinity (6,7,8). Here we studied the ability of wild type and mutant NCs purified from bacteria to bind zinc *in vitro*. According to the blotting technique reported by Schiff *et al.* (36), NCp10 was either transferred from a polyacrylamide gel (fig. 3A and B) or directly spotted (fig. 3C) onto nitrocellulose membrane. NCp10 was then probed with radioactive $^{65}ZnCl_2$. In these conditions, the *E.coli*-purified NCp10 showed an affinity for zinc comparable to that of viral NCp10 (fig. 3A and B, lanes 3 and 4, and fig. 3C). Furthermore, the affinity for zinc of lysine and tryptophane-mutated NCp10

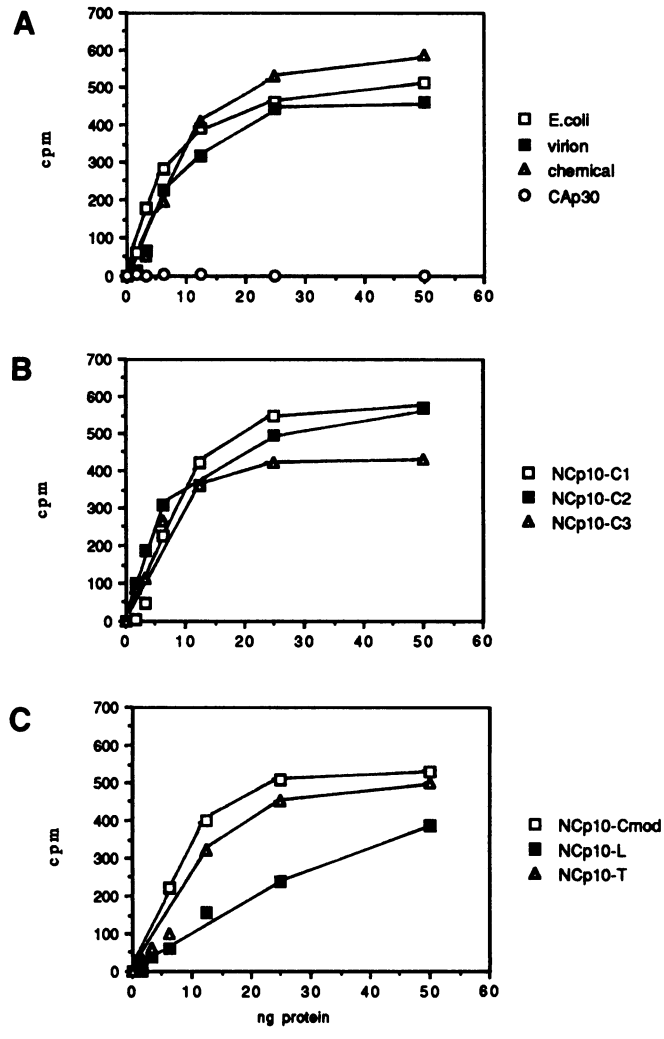

Figure 4. Affinity of wild type and mutant NCp10 for MuLV *Psi* RNA. Wild type and modified NCp10 were assayed in RNA binding assays on dot blots (Materials and Methods). The capsid protein of MoMuLV, CAp30, was used as a negative control. Increasing amounts of each protein species were spotted onto nitrocellulose membrane, and the membrane was incubated with a ^{32}P-labelled RNA containing nt 215 to 420 of MoMuLV genome at a 10 ng/ml concentration, in presence of 10 μg/ml of tRNA. MuLV RNA was synthesized *in vitro* using T7 RNA polymerase and purified as described before (27). The amount of ^{32}P-RNA bound to each spot of NCp10 was counted and the cpm were reported on a diagram as a function of NCp10 amount (ng). A) Control CAp30 and wild type NCp10 from virions, bacteria or chemical synthesis. B) Cys-mutant NCp10-C1, C2 and C3. C) Thiol alkylated NCp10 C^m, Trp-mutant NCp10-T and Lys-mutant NCp10-L. The protein species corresponding to each diagram is indicated on the right.

(referred to as NCp10-L and -T, respectively) was not altered (fig.3C), whereas NC mutated for one, two or three cysteine residues (NCp10-C1, -C2 and -C3) did not bind the metal ion under these conditions (fig.3A and B, lanes 1 and 2, and fig.3C). These results confirmed that the cysteine residues of the finger are required for zinc binding (8) whereas the tryptophane, and the lysine residues are not.

Binding of MuLV RNA to wild type, mutant and thiol-alkylated NCp10 proteins

To know whether the affinity of NCp10 for MuLV RNA is dependent on the finger domain and on the coordination of a zinc ion, we analyzed the RNA binding activity of the various mutant NCp10 proteins. The role of the zinc finger in RNA binding was also established by using a synthetic NCp10 protein (8). The cysteine residues of this chemical NCp10 were blocked by thiol alkylation (Material and Methods), making the protein unable to bind zinc (this modified NC is referred to as NCp10-C^m).

The effect of the different NC mutations or modification on RNA binding was studied by a dot blot assay. Increasing amounts of each NCp10 protein were spotted onto nitrocellulose membrane and probed with a ^{32}P-labelled MuLV RNA containing nucleotides 215 to 420 (the downstream part of *Psi* element between nt 400 and 567 is not required for encapsidation of MoMuLV genomic RNA, réf. 39), in presence of saturating amounts of non specific *E.coli* tRNA. The wild type NCp10 from *E.coli* was found to bind RNA as well as the wild type NCp10 purified from MuLV virions or chemically synthesized (fig. 4A). Furthermore all NC variants showed an affinity for MuLV RNA similar to that of wild type NCp10, except for NCp10-L which had a lower affinity (fig. 4 B and C). This indicated that under these experimental conditions the lysine residues next to the zinc finger are involved in RNA binding, whereas the cysteine and tryptophane residues located within the finger probably are not.

RNA annealing activity of wild type, mutant and thiol alkylated NCp10 proteins

In previous reports, we have shown that retroviral NC protein facilitates two nucleic acid annealing reactions, that are viral *Psi*$^+$ RNA dimerization (27) and annealing of replication primer tRNA onto genomic RNA (23). To know whether these NC properties are dependent upon the coordination of zinc, we analyzed the RNA annealing activity of the wild type or modified NCp10 proteins.

RNA dimerization and $tRNA^{Pro}$ annealing were analyzed in the same assay (see Materials and Methods). A 0.6 kilobase-length RNA containing the 5′ part of MuLV genome (nt −30 to 567) with the primer binding site and the dimerization sequence was synthesized *in vitro* (27). RNA was heat denatured and incubated with 5′^{32}P-labelled $tRNA^{Pro}$ and increasing amounts of wild type or modified NCp10. After incubation, NCp10 was eliminated by phenol extraction and RNA analyzed by agarose gel electrophoresis in native conditions. RNA dimerization was detected by ethidium bromid staining (fig.5A), and ^{32}P-tRNA hybridization by autoradiography (fig.5B).

First the ability of NCp10 purified from *E.coli* to activate RNA dimerization and primer tRNA hybridization was compared to that of viral or synthetic NCp10. As shown in figure 5 and 6A, comparable levels of dimeric RNA and bound tRNAPro were obtained with increasing amounts of NCp10 either from virions, *E.coli*, or *in vitro* synthesis.

Then the modified NCp10 proteins were assayed in the same conditions as the wild type protein. As shown in figure 6, NCp10-C1, -C2, -C^m and -T showed RNA dimerization and tRNA annealing activities comparable to that of the wild type proteins. Even NCp10-C3 had RNA annealing activities, although with a slightly lower initial slope for RNA dimerization. On the contrary, both annealing reactions were impaired when NCp10-L was used (Fig. $6A_3$ and B_3). These results indicated that the lysine residues, involved in RNA binding (fig. 4C), are also required for NCp10 annealing activity, whereas cysteine and tryptophane residues are not. It appeared from this study that the RNA annealing activity of NCp10 is related to its RNA binding activity but probably does not involve the binding of zinc.

DISCUSSION

The NCp10 protein of MoMuLV displays a zinc finger of the form Cys-X2-Cys-X4-His-X4-Cys, required for viral RNA

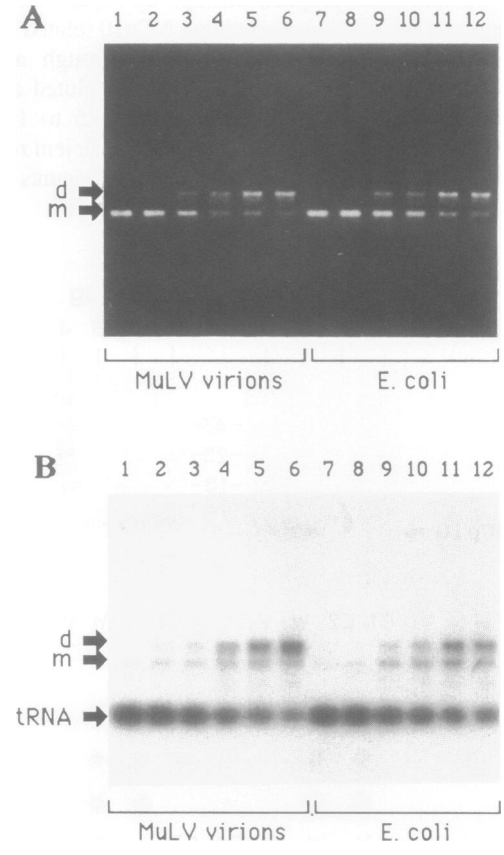

Figure 5. RNA annealing activities of virion and *E.coli* NCp10 proteins. NCp10 purified from MuLV virions or from *E.coli* was assayed for MuLV RNA dimerization and tRNA annealing onto MuLV RNA. As described in Materials and Methods, a 0.6 kb length-MoMuLV RNA (300ng) containing the dimerization sequence and the primer binding site was heat denatured at 100°C for 2 min, and incubated for 5 min with ^{32}P-labelled tRNAPro and increasing amounts of purified NCp10. RNA was then analyzed by electrophoresis on agarose gel in native conditions. A) RNA dimerization assay: an ethidium bromid staining of the gel is shown. RNA dimer and monomer migrations are indicated by m and d. Lanes 1−6: NCp10 from virions. Lanes 7−12: NCp10 from *E.coli*. From 1 to 6 and from 7 to 12, samples contained respectively 0, 12.5, 25, 50, 100 and 200 ng of NCp10. B) $tRNA^{Pro}$ annealing assay: The gel shown in A) was fixed by TCA 5%, dried and autoradiographed to detect tRNA. Free tRNA is indicated, as well as the positions of RNA monomer and dimer.

packaging *in vivo* (10,11). Here we describe modifications of the zinc finger of MoMuLV NCp10 by replacement or alkylation of aminoacid residues involved in zinc coordination. To analyze the RNA binding and annealing activities of the modified NCp10 proteins we have used MoMuLV RNA containing the primer binding site and the *Psi* element, that allow the specific annealing of the replication primer tRNA Pro and the selective dimerization and encapsidation of MuLV RNA, respectively (16,23,27). Previous *in vitro* studies have shown that cysteine modification of the zinc finger domain of NCp10 does not affect its affinity for polyribonucleotides (6,40,41). However the biological relevance of these results remains questionable since non viral, synthetic RNAs have been used.

Here we show that replacement or alkylation of the cysteine residues resulting in the loss of zinc binding to NCp10 does not alter the binding of NCp10 to MuLV RNA (fig.4). Similarly annealing of primer tRNAPro to the PBS is not significantly altered (fig.6B). Dimerization of the viral RNA does not seem to be drastically altered, although replacement of the three cysteines by serines slightly retards the process (fig. 6A). These results indicate that the viral RNA binding and annealing activities of NCp10 do not require zinc coordination under these experimental conditions.

Replacement of the tryptophane by a glycine in P10-T mutant is also without effect on the NCp10 dependent RNA binding and annealing processes. Previous studies (41) suggested an involvement of the tryptophane residue in NCp10-RNA interactions. Our *in vitro* results clearly show that even if that interaction exists, it is not critical for the NCp10 activities.

Two lysine residues located next to the zinc finger are highly conserved amongst retroviral NC proteins. As shown by Fu *et al.* (42), mutation of these two lysines in avian retroviral protein NCp12 disrupts its high affinity for the viral RNA and abolishes virus infectivity. This interesting observation prompted us to replace the two lysines (Lys41 and 42) located next to NCp10 finger by non charged aminoacids (Thr41, Ile42). This resulted

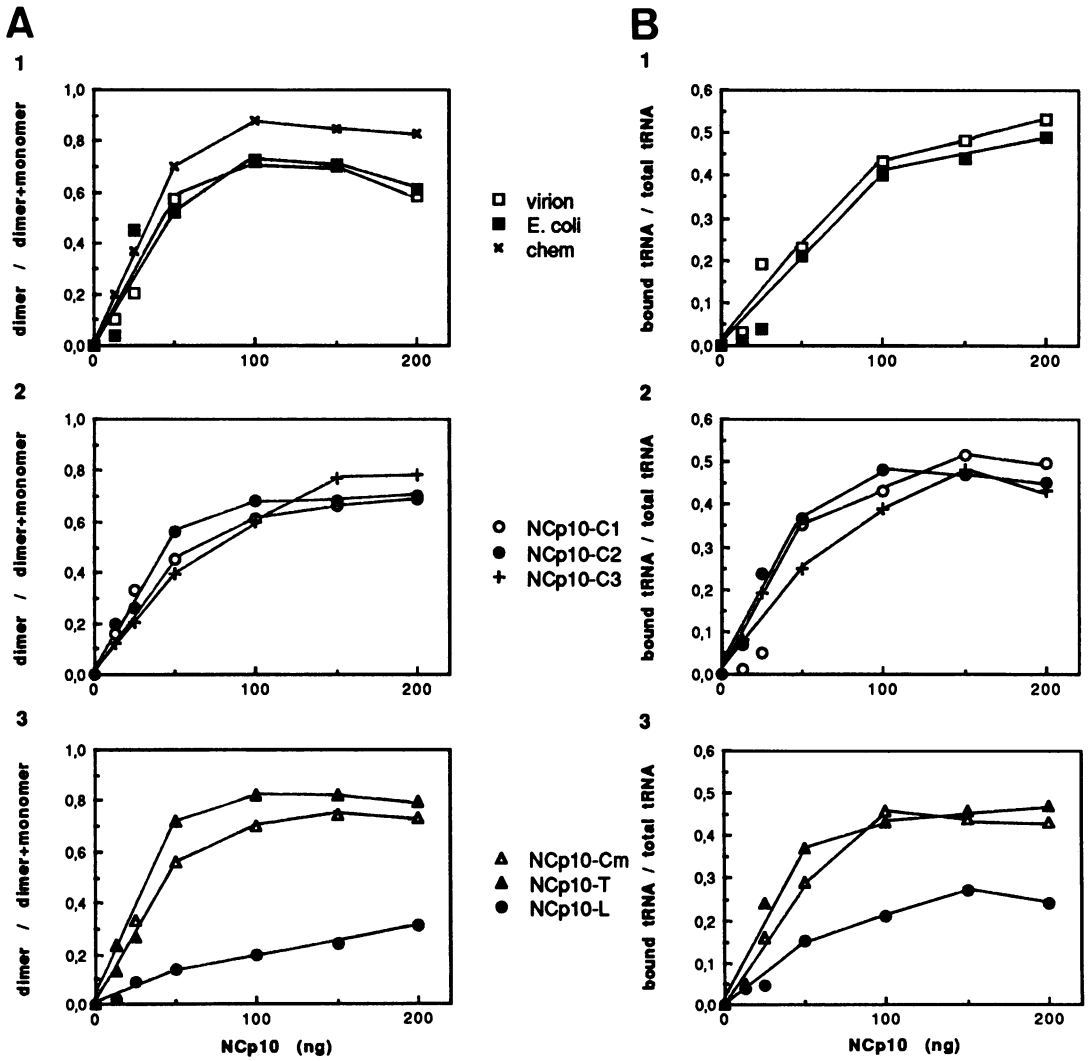

Figure 6. RNA annealing activities of modified NCp10. Wild type and modified NCp10 were assayed as in figure 5. Ratios of RNA dimerization and of $tRNA^{Pro}$ annealing were measured by scanning gel and autoradiogram, and reported on a diagram, as a function of NC protein amount (indicated in ng). Each diagram corresponds to an average of several experiments. A) *RNA dimerization assays*: ratios or MuLV RNA dimers towards total RNA are shown. 1/ Virion, E.coli and chemical wild type NCp10 ; 2/ Cys-mutant NCp10-C1, C2 and C3 ; 3/ thiol alkylated NCp10-C^m, Trp-mutant NCp10-T and Lys-mutant NCp10-L. B) *tRNA annealing assays*: ratios of tRNAPro annealed onto MuLV RNA towards total $tRNA^{Pro}$ are shown. 1/ Virion and *E.coli* wild type NCp10 ; 2/ Cys-mutant NCp10-C1, C2 and C3; 3/ thiol alkylated NCp10-C^m, Trp-mutant NCp10-T and Lys-mutant NCp10-L. The NCp10 species corresponding to each diagram is indicated on the right.

in a decrease of NCp10 RNA binding and annealing activities that are about three to more then ten-fold for viral RNA binding and dimerization, respectively (based on the amount of NCp10 necessary for 50% activity,fig. 4 and 6). This suggests that ionic interactions expected to take place between the lysines and the ribose-phosphate backbone of RNA are more crucial for the MoMuLV RNA dimerization process than for RNA binding. These findings also suggest that an optimal viral RNA annealing activity of the retroviral NC is required for virus infectivity.

Altogether these results seem to indicate that both annealing reactions depend only on the affinity of NCp10 for RNA and not on the zinc finger structure. Thus the specificity of the viral RNA annealing reactions would mostly be controlled by *cis* RNA sequences, the dimerization element, the primer binding site and the 3′ end of primer tRNA. This hypothesis is supported by the fact that these reactions can occur without NCp10 but in high salt and/or temperature conditions (27,43). Furthermore it has been observed that the retroviral NC protein activates the dimerization of heterologous viral RNA as well as the annealing of complementary DNA oligonucleotides onto RNA (23). In conclusion the retroviral NC protein appears to have a generalized nucleic acid annealing activity in physiological conditions, and this activity seems independent of zinc coordination.

Several types of finger domains with RNA and/or DNA binding properties and a strong affinity for Zn^{++} have been described, defining the Cys2-His2 finger proteins like transcription factor IIIA and the Cys2-Cys2 finger proteins like GAL4 and steroid hormone receptors (44). In these two cases, fingers have been shown to be directly involved in specific DNA (or RNA) recognition and binding (44). The NC protein, with its CCHC box, displays a somewhat different zinc finger structure, suggesting that it functions differently from the other proteins containing zinc finger domain(s). However, the zinc finger domain has been demonstrated as critically involved in genomic RNA packaging in virions (10,11). Among several explanations which could be proposed, one is suggested by comparision of the structure of NCp10 and of the gene 32 protein (g32P), a single-stranded DNA binding protein from bacteriophage T4. This protein contains a metal binding domain of the form Cys-X3-His-X5-Cys-X2-Cys which was found able to induce protein dimerization in complexed form, favoring cooperative binding of the protein to DNA (45). Such zinc finger dependent processes seem also to occur with the steroid hormone receptors in which one finger controls the specific DNA binding, while another finger might mediate the dimerization of receptor molecules (46). Therefore the zinc finger of NCp10 might favor self-association of this protein and/or its interaction with the capsid protein molecules during virion assembly.

Another hypothesis is suggested by the study of Crawford and Goff (47) showing that a deletion in MoMuLV protease, which blocks the processing of *gag* and *gag-pol* precursors, poorly impairs encapsidation of the genomic RNA. Thus RNA packaging could be controlled by an NC precursor rather than by the mature NCp10. Then the zinc finger might act by stabilizing a specific structure of this precursor, necessary for specific recognition of the *Psi* encapsidation element.

Nevertheless, the biochemical basis for the recognition of the viral Psi^+ RNA amongst the cellular RNAs remains an important question. As already pointed out, both the encapsidation element *Psi* and the retroviral NC domain are required for this recognition. In support of this assumption, preferential NC binding sites have been mapped in the encapsidation-dimerization element of avian and human genomic RNA (26,48). Then interactions between NC and the viral Psi^+ RNA would be the first step towards packaging of the retroviral genome, and the NC zinc finger with its coordinated zinc ion probably directs this selective recognition in the infected cells. Binding of the viral Psi^+ RNA to NC should result in RNA dimerization, cessation of RNA translation (25) and efficient RNA packaging. This is presently under investigation both *in vitro* and *in vivo* using biochemical and genetic approaches.

ACKNOWLEDGEMENTS

We thank G.Keith (Strasbourg, France) for a kind gift of $tRNA^{Pro}$ purified from beef liver. We thank C.Barat and P.Polard for helpful discussion, F.Amalric, J.-C.Faye, M.Guyader and A.Vincent for critical reading, M.Barbey for pictures and Y.de Préval for oligonucleotide synthesis. This work was supported by a grant from the Agence Nationale de Recherche sur le Sida (France).

REFERENCES

1. Bender, W., and Davidson, N. (1976) *Cell* **7**, 595–607.
2. Peters, G., Harad, F., Dahlberg, J.E., Parret, A., Haseltine, A. and Baltimore, D. (1977). *J. Virol.* **21**, 1031–1041.
3. Varmus, H. and Swanstrom, R. (1984) In R. Weiss, N. Teich, H. Varmus and J. Coffin (ed.) *RNA Tumor Viruses*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., Vol. I, pp. 369–512.
4. Fleissner, E. and Tress, E. (1973) *J. Virol.*, **12**, 1612–1615.
5. Méric, C., Darlix, J.L. and Spahr, P.F. (1984). *J. Mol. Biol.* **173**, 531–538.
6. Roberts, W.J., Pan, T., Elliott, J.I., Coleman, J.E. and Williams, K.R. (1989) *Biochemistry* **28**, 10043–10047.
7. Green, L.M. and Berg, J.M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6403–6407.
8. Cornille, F., Mely, Y., Ficheux, D., Savignol, I., Gerard, D., Darlix, J.L., Fournié-Zaluski, M.C. and Roques, B.P. (1990) *Int. J. Pept. Prot. Res.*, 36, 551–558.
9. Berg, J.M. (1986) *Science* 232, 485–487.
10. Gorelick, R.J., Henderson, L.E., Hanser, J.P. and Rein, A. (1988). *Proc. Natl. Acad. Sci. USA* **85**, 8420–8424.
11. Méric, C. and Goff, S.P. (1989). *J. Virol.* **63**, 1558–1568.
12. Méric, C. and Spahr, P.F. (1986). *J. Virol.* **60**, 450–459.
13. Dupraz, P., Oertle, S., Méric, C., Damay, P. and Spahr, P.F. (1990). *J. Virol.*, **64**, 4978–4987.
14. Aldovini, A. and Young, R.A. (1990) *J. Virol.* **64**, 1920–1926.
15. Watanabe, S. and Temin, H.M. (1982). *Proc. Natl. Acad. Sci. USA* 79, 5986–5990.
16. Mann, R., Mulligan, R.C. and Baltimore, D. (1983). *Cell* **33**, 153–159.
17. Pugatsch, C. and Stacey, D.W. (1983). *Virology*, **128**, 505–511.
18. Lever, A., Gottlinger, H., Haseltine, W. and Sodroski, J. (1989). *J. Virol.* **63**, 4085–4087.
19. Mann, R. and Baltimore, D. (1985). *J. Virol.* **54**, 401–407.
20. Davies, N.L., Scherer, M., Tsai, W.P. and Long, C. (1976). *J. Virol.*, **18**, 709–718.
21. Schulein, M., Burnette, W.N. and August, J.T. (1978) *J. Virol.* **26**, 54–60.
22. Nissen-Meyer, J.N. and Abraham, A.K. (1980) *J. Mol. Biol.* **142**, 19–28.
23. Prats, A.C., Sarih, L., Gabus, C., Litvak, S., Keith, G. and Darlix, J.L. (1988) *EMBO J.* **7**, 1777–1783.
24. Barat, C., Lullien, V., Schatz, O., Keith, G. and Darlix, J.L. (1989) *EMBO J.*, **8**, 3279–3285.
25. Bieth, E., Gabus, C. and Darlix, J.L. (1990). *Nucleic Acid Res.* **18**, 119–127.
26. Darlix, J.L., Gabus, C., Nugeyre, M.T., Clavel, F. and Barré-Sinoussi, F. (1990) *J. Mol. Biol.* 216, 689–699.
27. Prats, A.C., Roy, C., Wang, P., Erard, M., Housset, V., Gabus, C., Paoletti, C. and Darlix, J.L. (1990) *J. Virol.* **64**, 774–783.
28. Fan, H., and Baltimore, D. (1973) *J. Mol. Biol.* 28, 93–117.
29. Panganiban, A.T. and Fiore, D. (1988). *Science* **241**, 1064–1069.
30. Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) *Methods in Enzymology*, **185**, 60–89.
31. Shinnick, T.M., Lerner, R.A. and Sutcliffe, J.G. (1981) *Nature* **296**, 543–548.

32. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning : A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
33. Schägger, H., and von Jagow, G. (1987) *Anal. Biochem.* **166**, 368–379.
34. Strand, M. and August, J.T. (1976). *J. Biol. Chem.* **251**, 559–564.
35. Crestfield, A.M., Moore, S. and Stein, W.H. (1963). *J. Biol. Chem.* 238, 622–627.
36. Schiff, L.A., Nibert, M.L., and Fields, B.N. (1988), *Proc. Natl. Acad. Sci. USA* **85**, 4195–4199.
37. Bowen, B., Steinberg, J. , Laemmli, V.K., and Weintrong, H. (1980). *Nucleic Acid Res.* 8, 1–20.
38. Hemsley, A., Arnheim, A., Toney, M. D., Cortopassy, G., and Galas, D.J. (1989) *Nucleic Acid Res.*, 17, 6545–6551.
39. Schwarzberg, P., Colicelli, J. and Goff, S. (1983). *J. Virol.* **46**, 538–546.
40. Karpel, R.L., Henderson, L.E. and Oroszlan, S. (1987). *J. Biol. Chem.* **262**, 4961–4967.
41. Casas-Finet, J.R., Jhon, N.I. and Maki, A.H. (1988). *Biochemistry*, **27**, 1172–1178.
42. Fu, X., Katz, R.A., Skalka, A.M. and Leis, J. (1988) *J. Biol. Chem.* **263**, 2140–2145.
43. Roy, C., Tounekti, N., Mougel, M., Darlix, J.L., Paoletti, C., Ehresmann, C., Ehresmann, B. and Paoletti, J. (1990). *Nucleic Acid Res.*, 18, 7287–7292.
44. Berg., J.M. (1990) *J. Biol. Chem.* **265**, 6513–6516.
45. Pan, T., Giedroc, D.P. and Coleman, J.E. (1989). *Biochemistry*, **28**, 8828–8832.
46. Green, S., Kumar, V., Theulaz, I., Wahli, W. and Chambon, P. (1988). *EMBO J.*, **7**, 3037–3044.
47. Crawford, S. and Goff, S.P. (1985). *J. Virol.*, **53**, 899–907.
48. Darlix, J.L. and Spahr, P.F. (1982) *J. Mol. Biol.* 160, 147–161.