2434–2441 Nucleic Acids Research, 1995, Vol. 23, No. 13

Analysis of the nucleic acid annealing activities of nucleocapsid protein from HIV-1

Mary Lapadat-Tapolsky, Christine Pernelle[1], Christophe Borie[1] and Jean-Luc Darlix*

Unité de Virologie Humaine (INSERM 412), LaboRétro, Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69364 Lyon cedex 07, France and [1]Rhône-Poulenc Rorer, Research and Development, 13 Quai Jules Guesde, BP 14, 94403 Vitry-sur-Seine cedex, France

Received April 5, 1995; Revised and Accepted May 24, 1995

ABSTRACT

Retroviral nucleocapsid (NC) protein is an integral part of the virion nucleocapsid where it is in tight association with genomic RNA and the tRNA primer. NC protein is necessary for the dimerization and encapsidation of genomic RNA, the annealing of the tRNA primer to the primer binding site (PBS) and the initial strand transfer event. Due to the general nature of NC protein-promoted annealing, its use to improve nucleic acid interactions in various reactions can be envisioned. Parameters affecting NC-promoted nucleic acid annealing of NCp7 from HIV-1 have been analyzed. The promotion of RNA:RNA and RNA:DNA annealing by NCp7 is more sensitive to the concentration of $MgCl_2$ than the promotion of DNA:DNA hybridization. Stimulation of complex formation for all three complexes was efficient at 0–90 mM NaCl, between 23 and 55°C and at pH values between 6.5 and 9.5, inclusive. Parameters affecting NCp7-promoted hybridization of $tRNA^{Lys,3}$ to the PBS, which appears to be specific for NC protein, will be discussed. Results implicate the basic regions of NCp7, but not the zinc fingers, in promoting the annealing of complementary nucleic acid sequences. Finally, NCp7 strand transfer activity aids the formation of the most stable nucleic acid complex.

INTRODUCTION

In mature HIV-1 virions, the nucleocapsid structure, which lies within the capsid, contains the genomic 70S RNA dimer and the $tRNA^{Lys,3}$ primer in tight association with nucleocapsid protein (NC) molecules (1). NC protein, a well-conserved, small basic protein coded for by the 3′ end of *gag* (2–3), also has functional activities. NC protein, either as part of the gag polyprotein precursor and/or mature, appears to be necessary for virion formation and NC protein has been shown to be necessary for the dimerization and encapsidation of genomic RNA (1,4–9). In addition, NC protein has been shown to possess nucleic acid annealing activities which are important during reverse transcription. NC protein stimulates the initial step of reverse transcription, the annealing of the tRNA primer to the primer binding site (10–11). NC protein also stimulates the initial strand transfer event, probably by promoting the annealing of the R region of the strong stop cDNA with the complementary region of the 3′ end of the genomic RNA (12–14).

Nucleocapsid protein must thus promote the annealing of complementary RNA sequences, complementary DNA sequences and complementary RNA:DNA complexes. Although potent nucleic acid annealing activity of NC protein has been demonstrated (1,10,12,15–18), an in depth analysis of the parameters affecting this activity has not been undertaken. We describe here the effects of parameter variation on NCp7-promoted DNA:DNA, DNA:RNA, RNA:RNA and tRNA:RNA formation. The majority of the annealing studies were performed using RNA and DNA substrates corresponding to the repeated R region of the HIV-1 genome. This sequence was chosen because (i) it represents sequences important during the initial strand transfer during reverse transcription and (ii) it contains the stable stem–loop structure TAR which needs to be opened up in order to bind to the complementary region of the 3′ end of the genomic RNA. The sequence TAR [56 nucleotides (nt)], whose structure is conserved, was also used in many studies as were $tRNA^{Lys,3}$ and an RNA corresponding to the 5′ end of the genome and containing the PBS. The effects of various parameters such as ionic conditions, pH, temperature and nucleic acid and protein concentrations were examined. NCp7 promotes the annealing of complementary nucleic acid strands under conditions which do not normally promote complex formation. NCp7-promoted annealing is most rapid at 37°C and requires levels of NCp7 which saturate the nucleic acid present. NCp7 functions efficiently over a wide range of temperatures, pH, and nucleic acid and NaCl concentrations. The promotion of RNA:RNA and RNA:DNA annealing by NCp7 appears to be more sensitive to the concentration of $MgCl_2$ present than the promotion of complementary DNA strand hybridization. Promoting the annealing of complementary nucleic acid sequences is a general activity of this protein. NCp7-promoted hybridization of $tRNA^{Lys,3}$ to the PBS appears, however, to be specific for NC protein.

NCp7 is a basic protein with two zinc fingers of the form $CX_2CX_4HX_4C$ (CCHC) flanked by regions rich in basic residues (2,19). Recent *in vitro* studies have demonstrated that the regions of basic amino acids but not the zinc fingers of NCp7 are important for promoting the dimerization of genomic RNA, the

* To whom correspondence should be addressed

annealing of the tRNA primer to the PBS, synthesis of strong stop cDNA and the stimulation of the initial strand transfer event during reverse transcription (13,20). Results indicate that the basic regions, but not the zinc fingers, are directly implicated in promoting the annealing of complementary nucleic acid sequences *in vitro*.

Based on the importance of NC protein during several stages of the viral replication cycle, inhibition of the activity of this protein could provide a way to prevent infection and/or the lethal effects of HIV-1 infection. Previous results indicated that NC protein promotes the annealing of complementary nucleic acid strands even in the presence of a large excess of non-homologous DNA (15) and NCp7-promoted nucleic acid annealing has been exploited to increase the specificity and rate of hammerhead ribozyme catalysis (16–18). The use of this activity in improving and/or optimizing nucleic acid interactions in various reactions can be envisioned. In addition, a screening assay designed to target products which may inhibit HIV-1 replication *in vivo* based on their ability to inhibit NCp7-promoted nucleic acid annealing can be envisioned.

MATERIALS AND METHODS

RNA and DNA substrates

Oligonucleotides corresponding to the region R and TAR of HIV-1 (pmal strain) and those used to generate templates for preparation of RNAs representing the sense and antisense HIV-1 R region were purchased from either Eurogentec or Genset or supplied by Rhône-Poulenc Rorer. Those used for nucleic acid annealing assays were purified by urea–PAGE before use. HIV-1 R oligos were 96 nt and corresponded to the R region with minor changes at the ends; the minor changes were necessary for the preparation of complementary RNAs.

HIV-1 R sense [R(+)]: (5′)GGTCTCTCTTGTTGACCAGGCGAGCCCGGGAGTCTCTGGCTAGCAAGGCCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCCTCCC(3′)

HIV-1 R antisense [R(–)]: (5′)GGGAGGCACTCAAGGCAGCTTTATTGAGGCTTAAGCAGTGGGTTCCTTGCTAGCCAGAGAGCTCCCGGGCTCGACCTGGTCTAACAAGAGAGACC(3′)

HIV-1 R antisense oligonucleotides in which 10 nt were mutated at the 3′ or 5′ end were prepared by chemical synthesis. The underlined nucleotides indicate the positions mutated:

R(–).modified3′: (5′)GGGAGGCACTCAAGGCAAGCTTTTTGAGGCTTAAGCAGTGGGTTCCTTGCTAGCCAGAGAGCTCCCGGGCTCGACCTGGTCTAACTCAGTCTCTA(3′)

R(–).modified5′: (5′)ACTGATACTCCAAGGCAACTTTATTGAGGCTTAAGCAGTGGGTTCCTTGCTAGCCAGAGAGCTCCCGGGCTCGACCTGGTCTAACAAGAGAGACC(3′)

HIV-1 TAR sense and antisense (56 nt) oligonucleotides represent, respectively, the sense and antisense strands of the TAR region.

HIV-1 TAR sense [TAR(+)]: (5′)GGTCTCTCTTGTTAGACCAGGTCGAGCCCGGGAGCTCTCTGGCTAGCAAGGAACCC(3′)

HIV-1 TAR antisense [TAR(–)]: (5′)GGGTTCCTTGCTAGCCAGAGAGCTCCCGGGCTCGACCTGGTCTAACAAGAGAGACC(3′)

HIV-1 R sense RNA [RNA(+)]: a fragment corresponding to the R sequence listed above and containing a *Pst*I site and a T7 polymerase promoter upstream and a *Pst*I site downstream of the sequence was generated by PCR and cloned into pSP64 digested with *Pst*I. *Pst*I linearized DNA was treated with Klenow and transcribed with T7 RNA polymerase to generate the 96 nt sense RNA.

HIV-1 R antisense RNA [RNA(–)]: a fragment corresponding to the R sequence listed above and containing a *Pst*I site upstream and a T3 polymerase promoter and a *Xba*I site downstream was generated by PCR and cloned into pSP64 digested with *Xba*I and *Pst*I. *Pst*I linearized DNA was treated with Klenow and transcribed with T3 RNA polymerase to generate the 96 nt antisense RNA.

$tRNA^{Lys,3}$ was prepared by T7 RNA polymerase transcription of pTL9 DNA digested with *Ban*I (21). 5′ RNA was prepared using the clone pJF4 provided by C. Roy (Rhône-Poulenc Rorer) in which a region of pHIVCG4 (1) containing the first 311 nt of the 5′ end of the HIV-1 genome was subcloned into the vector pSP64 polyA. *Eco*RI linearized DNA was transcribed with T7 RNA polymerase to generate a 341 nt RNA (5′ 311 nt of the HIV-1 genome with a 30 nt polyA tail) which contains the PBS sequence. Pan RNA was prepared as described previously (22) using the clone provided by F. Baudin. This molecule permits intramolecular interactions between complementary 5′ and 3′ ends.

ColE1 sense RNA: the fragment of pBR322 corresponding to the ColE1 region (23) was amplified by PCR using an oligonucleotide corresponding to the 5′ end of the sense strand with an *Eco*RI restriction site and the T7 RNA polymerase promoter and a second complementary to the 3′ end of the sense strand with *Nae*I and *Xba*I restriction sites. This fragment was digested with *Eco*RI and *Xba*I and cloned into pSP64 digested with the same restriction enzymes. ColE1 sense RNA (108 nt; see ref. 23 for sequence) was obtained by transcription of *Nae*I digested DNA with T7 RNA polymerase.

ColE1 antisense RNA: the fragment of pBR322 corresponding to the ColE1 region was amplified by PCR using an oligonucleotide corresponding to the 5′ end of the sense strand with a *Hin*cII restriction site and a second complementary to the 3′ end of the sense strand with an *Eco*RI restriction site and the T3 RNA polymerase promoter. This fragment was digested with *Eco*RI and *Hin*cII and cloned into pSP64 digested with the same restriction enzymes. ColE1 antisense RNA (108 nt) was obtained by transcription of *Hin*cII digested DNA with T3 RNA polymerase.

All radiolabelled DNAs were prepared using T4 polynucleotide kinase and [γ-^{32}P]ATP. Radiolabelled RNA was obtained by using [α-^{32}P]CTP and/or [α-^{32}P]UTP in the transcription assay. Radiolabelled nucleic acid was subsequently purified by urea–PAGE. Restriction enzymes were from Promega or Eurogentec, all chemicals were of the highest grade available and prefabricated SDS–PAGE gels were obtained from Novex. p53 was kindly provided by L. Debussche and M.-C. Multon (Rhône-Poulenc Rorer).

Preparation of NCp7 and NCp7 peptides

NCp7 (72 amino acids) and NCp7 peptides (Fig. 5A) were prepared by peptide synthesis as described previously (20,24). The NCp7 peptides used were: NCp7(29-72) (44 amino acids) which lacks the N-terminal 28 amino acids, NCp7(1-72-Zn) (47 amino acids) in which the two zinc fingers were removed and replaced by glycine–glycine linkages, NCp7(13-72) (35 amino acids) which lacks the N-terminal 12 amino acids and both zinc fingers, NCp7(13-64) (52 amino acids) which lacks the 12

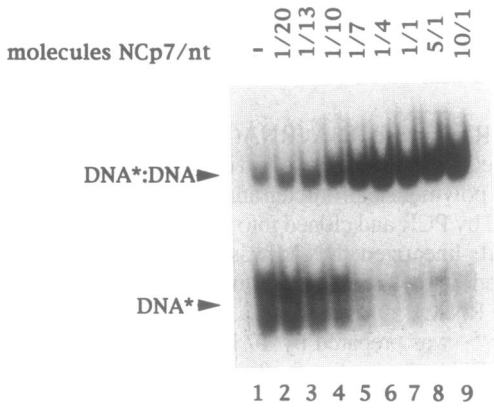

Figure 1. NCp7 promotes the annealing of complementary DNA sequences. HIV R(+) (15 fmol) and labelled HIV R(–) (15 fmol) were incubated as described in the Materials and Methods in the absence (lane 1) or presence of 0.14–57.5 pmol NCp7 such that the ratio of molecules of NCp7 per nucleotide indicated above each lane was present (lanes 2–9). NCp7 was subsequently removed from the reaction mixture by extraction with phenol–chloroform and the samples resolved by 6% native PAGE. DNA* indicates the migration of uncomplexed R(–) and DNA*:DNA represents the migration of the R(–):R(+) complex.

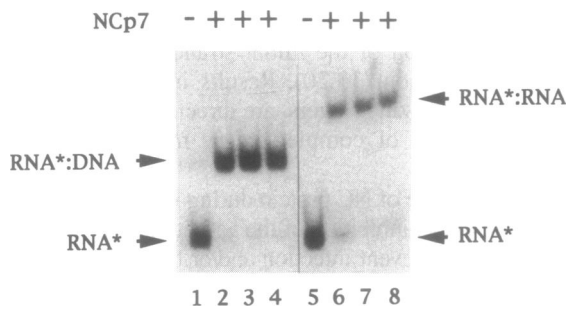

Figure 2. NCp7 promotes the formation of RNA:RNA and RNA:DNA complexes. Complementary RNA and DNA sequences (0.12 pmol each strand) (lanes 1–4) or complementary RNA sequences (0.12 pmol each strand) (lanes 5–8) representing the region R of HIV-1 were incubated in the absence (lanes 1 and 5) or presence of 23 pmol NCp7 (lanes 2–4 and 6–8) for 30 s (lanes 2 and 6), 2 min (lanes 3 and 7) or 5 min (lanes 1, 4, 5 and 8) and processed as described in the Materials and Methods. RNA* represents the migration of the radiolabelled uncomplexed RNA(–) while RNA*:DNA and RNA*:RNA represent the migration of the DNA(+):RNA(–) and RNA(+):RNA(–) complexes, respectively.

N-terminal and eight C-terminal amino acids, NCp7(13-56) (19 amino acids) which lacks the N-terminal 12 amino acids, both zinc fingers and the 16 C-terminal amino acids and NCp7(51-72) (22 amino acids) which lacks the N-terminal 50 amino acids.

Nucleic acid annealing assays

Complementary DNA and/or RNA oligonucleotides (0.03–0.12 pmol total), one of which was radiolabelled, were incubated in a final volume of 20–40 µl in buffer containing (unless indicated otherwise) 30 mM Tris-Cl (pH 7.5), 30 mM NaCl, 0.2 mM $MgCl_2$, 10 µM $ZnCl_2$ and 5 mM DTT in the presence or absence of NCp7 at 37°C. The reaction was stopped by the addition of SDS (to 1%) and the nucleic acid separated from the denatured protein by extraction with phenol–chloroform. The samples were resolved by native PAGE and the results viewed by autoradiography. The percentage of nucleic acid present in complex was determined by densitometric scanning. When testing the effect of $MgCl_2$ concentration, a concentration of salt equivalent to 60 mM NaCl was used per reaction.

For the annealing of $tRNA^{Lys,3}$ to 5′ RNA, 0.39 pmol $tRNA^{Lys,3}$ and 0.95 pmol 5′ RNA were preincubated for 10 min at 37°C in a final volume of 10 µl in buffer containing 25 mM Tris-Cl (pH 7.5), 100 mM NaCl, 0.8 mM $MgCl_2$ and 5 mM DTT. NCp7 (24 pmol) was then added and the reaction stopped as described above. NCp7 was digested with proteinase K (0.33 mg/ml final concentration; 10 min, 37°C) prior to removal by extraction with phenol–chloroform, the samples resolved by 10% SDS–PAGE and the results analyzed by phosphoimaging.

Annealing of complementary DNA sequences by p53 was performed in a final volume of 10 µl with 0.015 pmol of TAR(+) and TAR(–) in buffer containing 25 mM Tris-Cl (pH 7.5), 20 mM NaCl, 5 mM DTT and 1 mg/ml BSA for 10 min at 37°C. The reactions were stopped by the addition of SDS (0.33%) and proteinase K (0.33 mg/ml) during a 10 min incubation at 37°C and the nucleic acid extracted with phenol–chloroform and resolved by 10% SDS–PAGE. The results were quantified by phosphoimaging. Annealing of $tRNA^{Lys,3}$ to 5′ RNA by p53 was performed likewise with 0.39 pmol $tRNA^{Lys,3}$ and 0.95 pmol 5′ RNA in buffer containing 25 mM Tris-Cl (pH 7.5), 100 mM NaCl, 5 mM DTT and 10 µg/ml BSA.

RESULTS

NCp7 promotes annealing of complementary nucleic acid sequences

The ability of NCp7 to promote annealing of complementary DNA sequences with a high degree of secondary structure was examined by incubating complementary 56mers or 96mers corresponding to the region TAR or R, respectively, of HIV-1. NCp7 was able to stimulate the annealing of complemenatary 56mers (data not shown) as well as complementary 96mers (Fig. 1). In the absence of NCp7, only a very low level of spontaneous complex formation was observed after a 5 min incubation at 37°C (lane 1), indicating that the conditions used do not favor annealing of complementary strands. Incubation of the complementary strands at 37°C in the absence of NCp7 even for extended periods of time still did not result in the accumulation of significant levels of complex (data not shown). However, in the presence of increasing levels of NCp7, an increase in the level of annealed DNA was observed (lanes 2–9). These results also indicate that approximately one molecule of NCp7 per 4–7 nt of single-stranded DNA is required to optimally promote complex formation. This ratio corresponds to the NCp7 occlusion binding site (15,25–26) and indicates that the DNA must be covered for the NC protein to exert its effect, in agreement with previous work (15).

The product of the NCp7-promoted annealing reaction is a double-stranded nucleic acid; no large complexes were observed. In addition, as NCp7 was removed prior to analysis, the product is necessarily double-stranded and not two strands associated via protein–nucleic acid interactions. Finally, it should be noted that ATP is not required for this reaction, in agreement with previous work (27).

In addition to promoting the annealing of complementary DNA sequences, NCp7 also stimulates the annealing of complementary

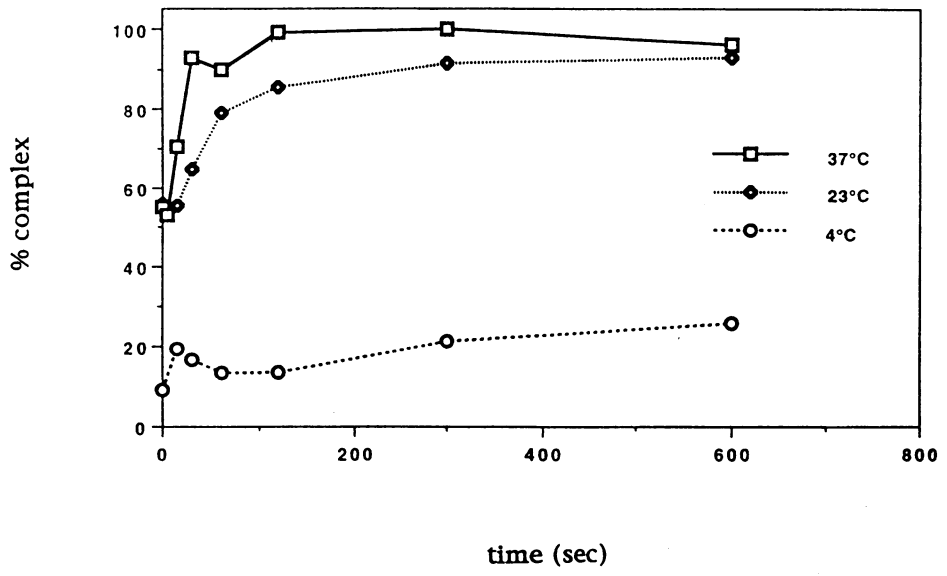

Figure 3. Temperature dependence of the rate of NCp7 promoted annealing of complementary DNA sequences. R(+) and R(–) (60 fmol each) were incubated in the presence of 2.9 pmol NCp7 at 4°C (circles), 23°C (ovals) or 37°C (squares) as described. The reactions were stopped at specific intervals and the percentage of DNA present as complex determined by densiometric scanning of the autoradiograph.

RNA and DNA sequences as well as complementary RNA sequences (Fig. 2). Incubation of complementary DNA and RNA sequences corresponding to the region R of HIV-1 for 5 min at 37°C in the absence of NCp7 resulted in a very low level of spontaneous complex formation (lanes 1 and 5). However, in the presence of saturating levels of NCp7 most of the nucleic acid was double-stranded (lanes 2–4 and 6–8). These results, and the fact that NCp7 promotes the annealing of complementary RNA sequences corresponding to RNA1 of ColE1, 108 nt highly structured RNAs (23) (data not shown), indicate that the nucleic acid annealing activity of NCp7 is a general activity of this protein.

Factors affecting the nucleic acid annealing activity of NCp7

The effect of various parameters on the ability of NCp7 to promote the annealing of complementary DNA, complementary RNA and complementary DNA and RNA sequences has been examined. In general, the effects were similar for all three complexes, but some differences were observed.

The nucleic acid annealing activity of NCp7 functions most efficiently at 37°C for RNA:RNA, RNA:DNA and DNA:DNA complex formation. An example of NCp7-promoted DNA:DNA formation at 4, 23 and 37°C is shown in Figure 3. An appreciable level of duplex DNA was observed after 15 s at 37°C and the reaction was basically complete after 2 min in the presence of saturating levels of NCp7. These results are consistent with those obtained previously for the renaturation of DNA complexes (15,27) and confirm that NCp7-promoted hybridization is potent. NCp7 functions very inefficiently at 4°C as indicated by the low level of annealed complex observed after a 10 min incubation in the presence of saturating levels of NCp7. In addition, although the same maximum level of annealing was observed after 10 min at 23°C, the rate of promotion of complex formation was more rapid at 37°C. NCp7 begins to lose its ability to promote annealing of complementary DNA sequences at temperatures >55°C; between 23 and 45°C similar levels of complex

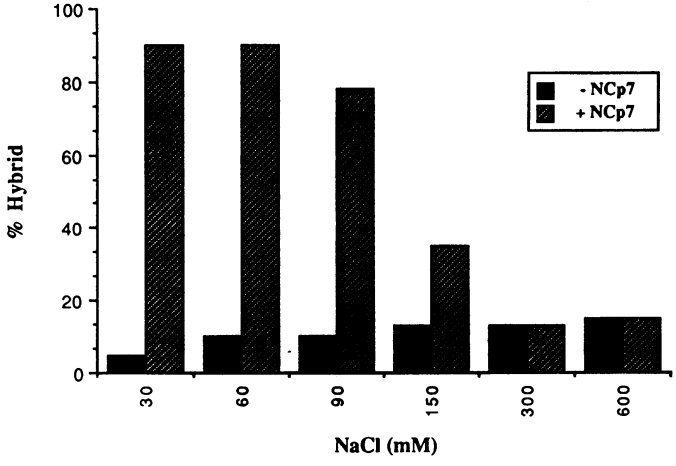

Figure 4. Effect of NaCl concentration on the ability of NCp7 to promote the annealing of complementary DNA sequences. R(+) and R(–) (60 fmol each) were incubated without (solid bars) or with (hatched bars) NCp7 (2.9 pmol) in the presence of varying concentrations of NaCl for 10 min at 37°C and the reactions processed as described in the Materials and Methods. The percentage of radiolabelled DNA present in the hybrid was determined by densitometric scanning of the autoradiograph.

formation were observed (data not shown). Similar results were obtained for NCp7-promoted DNA:RNA and RNA:RNA annealing (data not shown).

NCp7 does not require the presence of NaCl as a high level of annealing was observed at 0 mM NaCl (data not shown). NCp7 also efficiently stimulated the annealing of all three complexes at 30, 60 and 90 mM NaCl with a maximum of annealing observed at 30–60 mM NaCl. Figure 4 illustrates typical results obtained for all three complexes. A decrease in NCp7-promoted annealing was observed beginning at 90 mM NaCl and at >300 mM NaCl very little if any stimulation of annealing was seen.

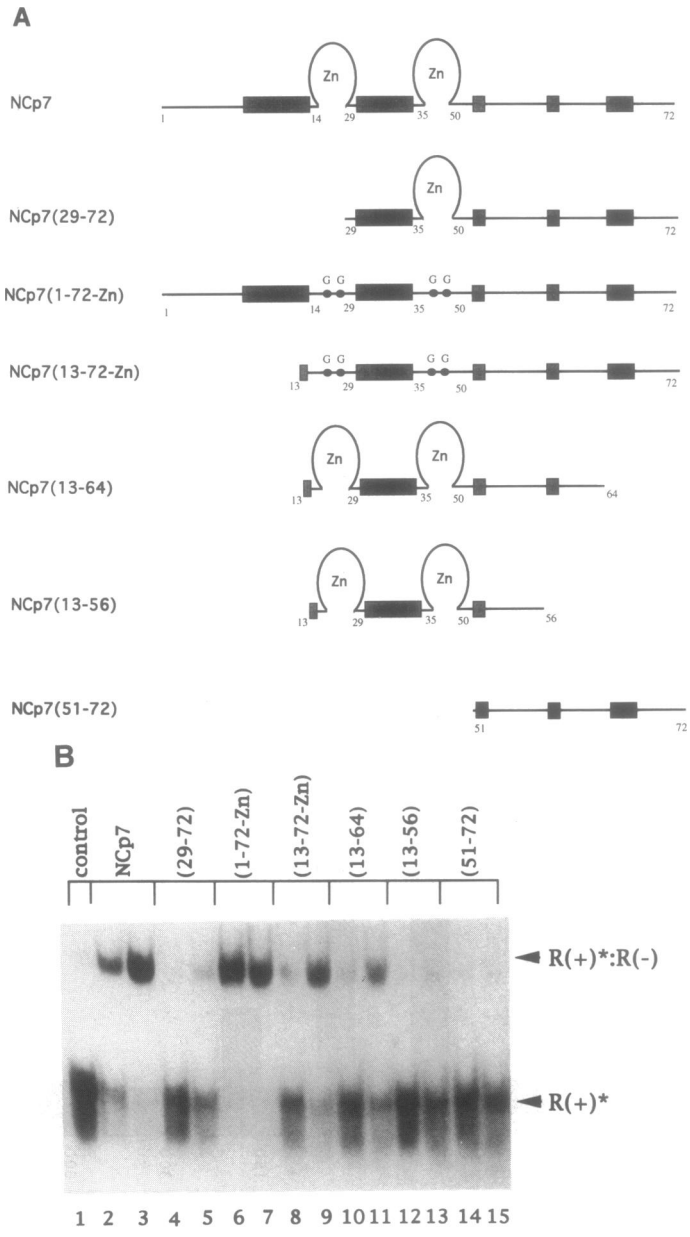

Figure 5. Ability of NCp7 peptides to promote annealing of complementary DNA sequences. (**A**) Scheme representing the structure of HIV-1 NCp7 and NCp7 peptides. The boxes represent the regions rich in basic amino acids. In the peptides NCp7(1-72-Zn) and NCp7(13-72-Zn), the zinc fingers were replaced by two glycine residues. The glycine–glycine linkage maintains the distance between the regions rich in basic amino acids observed in wild-type NCp7. (**B**) Complementary DNA strands representing the region R (30 fmol each) were incubated in the absence (control, lane 1) or presence of NCp7 (lanes 2 and 3) or an NCp7 peptide (lanes 4–15) as indicated at the top of the gel for 5 min at 37°C and the reactions processed as described in the Materials and Methods. Levels of NCp7 or NCp7 peptides used are as follows. NCp7: 1.2 pmol, lane 2; 2.9 pmol, lane 3. NCp7 peptides: 2.9 pmol, lanes 4–15 odd; 58 pmol, lanes 4–15 even. R(+)* indicates the migration of the radiolabelled uncomplexed sense strand and R(+)*:R(–) indicates the migration of the sense:antisense complex.

The effect of pH on the ability of NCp7 to promote annealing of complementary strands was examined. NCp7 was added to a reaction mixture which had been preincubated for 5 min at 37°C to allow the final pH value to stabilize. Results indicate that NCp7 functions efficiently at a pH of 6.5–9.5, with a slight maximum at pH 7.0, to promote annealing of all three nucleic acid complexes (data not shown).

The stimulation of annealing of DNA:DNA, RNA:DNA and RNA:RNA complexes by NCp7 appeared to be most sensitive to the concentration of $MgCl_2$. While NCp7 promoted DNA:DNA annealing efficiently over a range of $MgCl_2$ concentrations from 0 to 20 mM, a progressive 2–5-fold decrease in the ability of NCp7 to promote RNA:RNA and RNA:DNA annealing was observed at concentrations of $MgCl_2$ >1–2 mM (data not shown).

The ability of NCp7 peptides lacking either one or both zinc fingers and/or regions of basic amino acids (Fig. 5A) to promote the annealing of complementary DNA sequences was examined (Fig. 5B). The level of spontaneous annealing after a 10 min incubation of the complementary DNA sequences corresponding to the region R is seen in lane 1. As expected, in the presence of saturating levels of NCp7 most of the DNA was annealed to its complementary strand (lanes 2 and 3). However, in the presence of similar levels of peptides (lanes 4–15, even), only NCp7(1-72-Zn) (lane 7) efficiently promoted strand annealing. When the peptides were present in a 20-fold excess (lanes 4–15, odd), NCp7(13-72-Zn) (lane 9) and NCp7(13-64) (lane 11) were able to stimulate annealing of the complementary DNA strands while only background levels of spontaneous annealing were observed in the presence of the other peptides. These results indicate that, *in vitro*, the basic amino acids but not the zinc fingers are important for strand annealing.

Annealing of $tRNA^{Lys,3}$ to the PBS

As shown previously, NCp7 stimulates the annealing of *in vitro* synthesized $tRNA^{Lys,3}$ to an RNA representing the 5′ end of the HIV genome and containing the PBS (5′ RNA) (10). It is interesting to note that saturating levels of NCp7 were not required to observe a stimulation of tRNA:5′ RNA annealing by NCp7 as was seen for the annealing of complementary RNA and DNA sequences (data not shown).

NCp7-promoted annealing of $tRNA^{Lys,3}$ to 5′ RNA is also quite rapid; >80% $tRNA^{Lys,3}$:5′ RNA complex was observed after 1 min (Fig. 6). As observed for the annealing of complementary DNA sequences, maximum levels of $tRNA^{Lys,3}$ annealing to 5′ RNA were obtained in the presence of 30 mM NaCl; at >80 mM NaCl a significant decrease in hybridization was observed (data not shown). The effect of $MgCl_2$ concentration on NCp7-promoted annealing of $tRNA^{Lys,3}$ to 5′ RNA differed from that observed for RNA:RNA annealing; concentrations of $MgCl_2$ >2 mM reduced the maximum level of NCp7-promoted $tRNA^{Lys,3}$:5′RNA formation by only 25% (data not shown).

P53 has been shown to promote the annealing of complementary DNA and RNA sequences (28–29). Incubation of p53 in the presence of complementary DNA sequences representing the region TAR resulted in the formation of a stable hybrid (Fig. 7). However, p53 was unable to stimulate hybridization of $tRNA^{Lys,3}$ to 5′ RNA containing the PBS.

NCp7 promotes the formation of the most stable complex

NCp7 promotes the annealing of incompletely complementary DNA strands such as R(+):TAR(–) and R(+) to an incompletely complementary R(–) strand in which 10 nt were mutated at either the 3′ [R(–).modified3′) or 5′ [R(–).modified5′] end (Fig. 8; see Materials and Methods for details). When labelled R(+) was

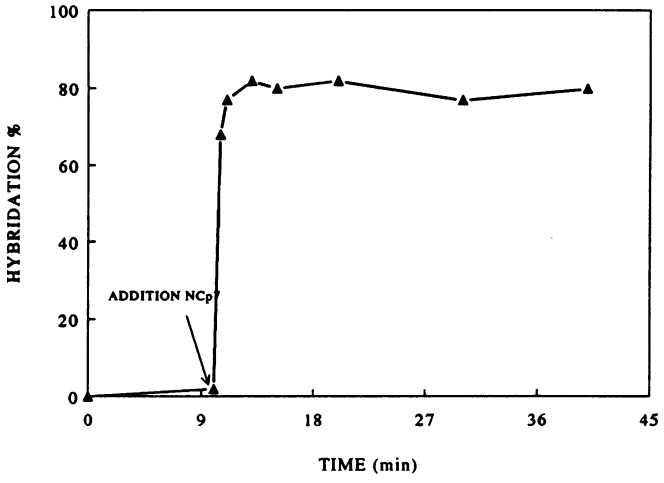

Figure 6. Kinetics of NCp7-promoted $tRNA^{Lys,3}$ annealing to 5′ RNA. $tRNA^{Lys,3}$ and 5′ RNA were preincubated for 10 min prior to the addition of NCp7 (see arrow) as described in the Materials and Methods. The reactions were stopped as a function of time and the percentage of labelled $tRNA^{Lys,3}$ complexed to 5′ RNA determined by phosphoimage analysis as described.

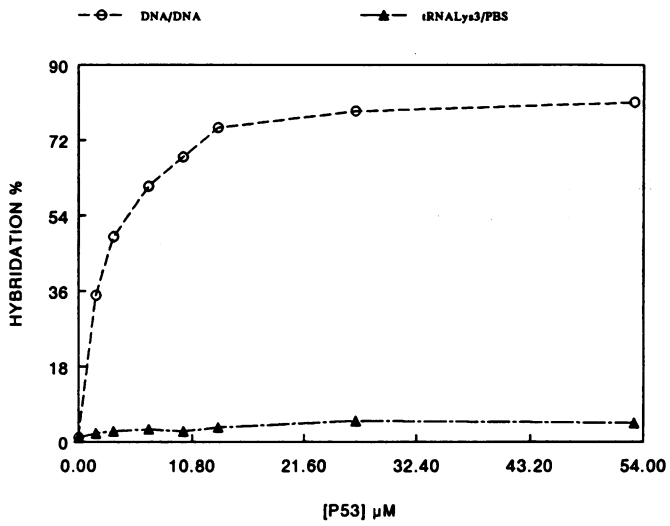

Figure 7. Ability of p53 to promote annealing of complementary DNA and RNA sequences. TAR(+) and TAR(–) (open circles) or $tRNA^{Lys,3}$ and 5′ RNA (filled triangles) were incubated in the absence or presence of various concentrations of p53 and the reactions processed as described in Materials and Methods. The level of duplex nucleic acid was determined by phosphoimage analysis.

incubated with R(–), R(–).modified3′ or R(–).modified5′ in the presence of NCp7, NCp7 promoted complex formation (lanes 2–4, respectively). Differential migration of R(+):R(–) and R(+):R(–).modified 3′ complexes permitted a comparison of the rates of formation of these two complexes. When equimolar levels of sense, antisense and R(–).modified3′ were incubated with NCp7, the preferential appearance of the sense:antisense complex was observed as a function of time (lanes 5–10). Equivalent levels of R(+):R(–) and R(+):R(–).modified3′ complexes were formed initially (15 s, lane 5). Then the effect of strand transfer was observed, leading to the preferential accumulation of the more stable complex, as defined by the number of base pairs formed, after 1.5 min [lanes 8–10, ~80% R(+):R(–)] (27). This result also indicates that NCp7-promoted annealing is a dynamic process.

DISCUSSION

The results presented here confirm that NCp7 from HIV-1 can promote the annealing of complementary nucleic acid sequences (15,27) and analyze in depth the factors affecting the strand annealing activity of NCp7 *in vitro*. The oligonucleotides used in this study represent the sequences R and TAR and are longer and more structured than oligonucleotides used in previous studies. In addition, the nucleic acid sequences employed represent the viral sequences important during the initiation of reverse transcription and the initial strand transfer event. As NCp7 was removed prior to analysis of the nucleic acid products, the double-stranded reaction products are hybrids and not complementary strands held together by the nucleocapsid protein. While low levels of NCp7 sometimes stimulated a low level of annealing, in general saturating levels of protein were required for promotion of hybridization as described previously (15). This result is similar to that obtained with several DNA binding proteins such as *Escherichia coli* SSB and phage T4 gene 32 protein which also require saturating levels of protein for optimal promotion of hybridization (30–31).

Comparatively little is known about the way in which NC proteins interact with nucleic acids. A recent study on NC protein DNA strand annealing activity demonstrated that the annealing of complementary DNA strands was second order both in the absence and presence of NC protein (15), similar to that observed with helix destabilizing proteins. This result was taken to indicate that NC protein probably does not function by promoting molecular aggregation. In addition, this was taken to indicate that the binding of NCp7 should be, as was observed, stoichiometric and that NCp7 most likely functions by destabilizing intramolecular interactions and promoting intermolecular interactions. This hypothesis is supported by our initial results obtained using a pan RNA molecule which contains complementary 5′ and 3′ ends which can hybridize intramolecularly to form a 'panhandle' (22). If NCp7 unwinds its nucleic acid substrates, no intrastrand annealing should be observed. In the presence of NCp7 pan RNA dimers and not the panhandle RNA structure were obtained, indicating that NCp7 promotes intermolecular interactions between pan RNA molecules. While the presence of the panhandle obtained by intramolecular interactions in the presence of NCp7 could not be entirely excluded, the absence of pan RNA dimers obtained using conditions designed to permit interactions between complementary regions (see ref. 22 for details) supports the idea that NCp7 preferentially promotes intermolecular interactions. Thus, NCp7 may, like the adenovirus DNA-binding protein (32), unwind and hold the nucleic acid in an extended conformation (preventing the possibility that the complementary 5′ and 3′ ends of the pan RNA molecule would come into contact). This is in agreement with results by Khan and Giedroc who illustrated the inability of Mg^{2+} to catalyze intramolecular base-pairing of tRNA in the presence of NCp7 while the mammalian helix destabilizing protein UP1 permitted the tRNA to re-fold fully under the same conditions (33).

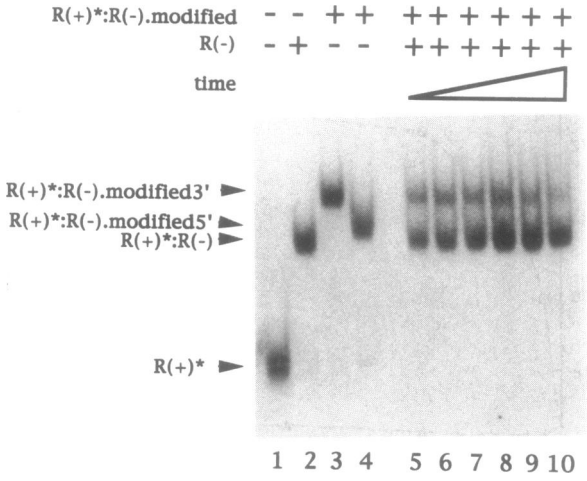

Figure 8. NCp7 promotes the annealing of the most stable complex. Radiolabelled R(+) was incubated alone (lane 1) or in the presence of R(–) (lane 2), R(–).modified3′ (lane 3) or R(–).modified5′ (lane 4) with NCp7 (1 molecule/4 nt) for 5 min as described. Alternatively, R(+) was incubated with equimolar levels of R(–) and R(–).modified3′ in the presence of NCp7 (lanes 5–10) for 15, 30, 60, 90, 120 or 300 s, respectively, and the relative amounts of sense:antisense and sense:antisense.modified complexes formed as a function of incubation time was determined by laser densitometric scanning.

The general ability of NCp7 to facilitate annealing of complementary nucleic acid sequences may also indicate that NCp7 functions by increasing the effective concentration of single-stranded molecules, a process known as molecular crowding (34). In addition, it is possible that as NC protein completely covers the DNA strands and functions to melt out secondary structures it increases the possibility that a nucleation event will be productive. Additional studies to determine the exact mechanism of the NCp7-promoted annealing reaction are necessary.

While NCp7 appears to function in many aspects similar to SSBs, its ability to stimulate the annealing of complementary RNA strands is different from the majority of these proteins as this does not appear to be a general property of SSBs. In addition, the ability to hybridize tRNA to RNA containing the PBS appears to be specific for NC protein as *E.coli* recA and phage T4 gene 32 protein were unable to promote this hybridization (10). Even p53, which promotes both DNA and RNA annealing as well as strand transfer (28–29), is unable to promote the binding of tRNA to the PBS (Fig. 7). The uniqueness of NCp7's flexibility could reflect the importance of the nucleic acid annealing activities of NC protein during reverse transcription which requires the formation of RNA:RNA (binding of primer to PBS), RNA:DNA (initial strand transfer event) and DNA:DNA complexes (second strand transfer event) as well as during forced or copy choice recombination events (35).

It has been published that NCp7 has the ability to unwind nucleic acids (27,29,36). Under the reaction conditions used, we were unable to detect de-annealing of nucleic acid complexes in the absence of an acceptor strand. It is possible that conditions for de-annealing may be different from those which promote annealing (27) and that our conditions do not permit efficient de-annealing. Or, NCp7 may, like the mammalian heterogenous nuclear ribonucleoprotein A1, facilitate the pairing of complementary nucleic acid strands by facilitated annealing rather than by using an unwinding mechanism (37). It is also possible that the formation of certain nucleic acid complexes (like tRNA annealing to an RNA containing the PBS) requires facilitated unwinding by NC protein while the formation of other nucleic acid complexes uses a different mechanism. This might explain why certain annealing reactions (such as the annealing of tRNA to the PBS) appear to be quite specific for NC protein even though it appears to have a general nucleic acid annealing activity.

We have shown that NCp7 promotes the annealing of various complementary nucleic acid sequences under a wide range of ionic, pH and temperature conditions. Only the effect of $MgCl_2$ concentration appeared to vary significantly for the various complexes. The fact that NCp7 must function both as part of the gag polyprotein and as a mature protein, within the cellular cytoplasm and within the virion nucleocapsid and that it interacts with DNA, RNA and tRNA during virion formation and reverse transcription may explain its activity under this wide range of conditions. Consequently, this activity may be relatively easily exploited for use in other systems where the annealing of complementary nucleic acid sequences is rate limiting. NCp7 has already been shown to increase the specificity and rate of hammerhead ribozyme catalysis (16–18). The nucleic acid annealing activities of NCp7 may potentially be used to improve the efficiency of site-directed mutagenesis or enhance the specificity of PCR amplification. In addition, as indicated by Tsuchihasi and Brown (27), other possible practical applications include the enhancement of DNA annealing to a DNA immobilized on a chip or other surface or improvement of the genomic mismatch scanning technique for genetic linkage mapping. Finally, based on the importance of NC protein during several stages of the viral life cycle, inhibition of the nucleic acid annealing activity of NC protein could provide a way to prevent HIV-1 replication. The nucleic acid annealing activity of NCp7 provides a simple assay by which to screen for compounds which may inhibit, *in vivo*, NCp7 activity.

ACKNOWLEDGEMENTS

We thank B. Roques and H. de Rocquigny for NCp7 and NCp7 peptides, F. Baudin for the pan RNA clone, L. Debussche and M.-C. Multon for p53, C. Roy for the 5′ RNA clone and J.-F. Mayaux for helpful discussions. This work was supported by grants from the French program against AIDS (ANRS) to J.-L.D. M.L.-T. acknowledges support by Rhône-Poulenc Rorer.

REFERENCES

1 Darlix,J.-L., Gabus,C., Nugeyre,M.-T., Clavel,F. and Barré-Sinoussi,F. (1990) *J. Mol. Biol.*, **216**, 689–699.
2 Wain-Hobson,S., Sonigo,P., Danos,O., Cole,S. and Alizon,M. (1985) *Cell*, **40**, 9–17.
3 Barré-Sinoussi,F., Chermann,J.C., Rey,F., Nugeyere,M.-T., Chamaret,S., Gruest,J., Danguet,C., Axler-Blin,C., Vézinet-Brun,F., Rouzioux,C., Rozenbaum,W. and Montagnier,L. (1983) *Science*, **220**, 868–870.
4 Bieth,E., Gabus,C. and Darlix,J.-L. (1990) *Nucleic Acids Res.*, **18**, 119–127.
5 Méric,C. and Spahr,P.-F. (1986) *J. Virol.*, **60**, 450–459.
6 Aldovini,A. and Young,R. (1990) *J. Virol.*, **64**, 1920–1926.
7 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.
8 Gottlinger,H., Sodroski,J. and Haseltine,W. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 5781–5785.
9 Méric,C. and Goff,S.P. (1989) *J. Virol.*, **63**, 1558–1568.
10 Prats,A.-C., Sarih,L., Gabus,C., Litvak,S., Keith,G. and Darlix,J.-L. (1988) *EMBO J.*, **7**, 1777–1783.

11 Barat,C., Lullien,V., Schatz,O., Keith,G., Nugeyere,M.-T., Gruninger-Leitch,F., Barré-Sinoussi,F., le Grice,S.F.J. and Darlix,J.-L. (1989) *EMBO J.*, **8**, 3279–3285.
12 Lapadat-Tapolsky,M., de Rocquigny,H., van Gent,D., Roques,B., Plasterk,R. and Darlix,J.-L. (1993) *Nucleic Acids Res.*, **21**, 831–839.
13 Darlix,J.-L., Vincent,A., Gabus,C., de Rocquigny,H. and Roques,B. (1993) *Comptes Rendus Acad. Sci.*, **316**, 763–771.
14 Allain,B., Lapadat-Tapolsky,M., Berlioz,C. and Darlix,J.-L. (1994) *EMBO J.*, **13**, 973–981.
15 Dib-Hajj,F., Khan,R. and Giedroc,D.P. (1993) *Protein Science*, **2**, 231–243.
16 Bertrand,E.L. and Rossi,J.J. (1994) *EMBO J.*, **13**, 2904–2912.
17 Tsuchihashi,Z., Khosla,M. and Herschlag,D. (1993) *Science*, **262**, 99–102.
18 Herschlag,D., Khosla,M., Tsuchihashi,Z. and Karpel,R.L. (1994) *EMBO J.*, **13**, 2913–2924.
19 Di Marzo-Veronese,F., Rahman,R., Copeland,T., Orsozlan,S., Gallo,R.C. and Sarnagadharan,M.G. (1987) *AIDS Res. Human Retroviruses*, **3**, 253–264.
20 De Rocquigny,H., Gabus,C., Vincent,A., Fournié-Zaluski,M.-C., Roques,B. and Darlix,J.-L. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 6472–6476.
21 Barat,C., Le Grice,S.F.J. and Darlix,J.-L. (1991) *Nucleic Acids Res.*, **19**, 51–757.
22 Baudin,F., Bach,C., Cusack,S. and Ruigrok,W.H. (1994) *EMBO J.*, **13**, 3158–3165.
23 Cesareni,G., Muesing,M. and Polisky,B. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 6313–6317.
24 De Rocquigny,H., Ficheux,D., Gabus,C., Fournié-Zaluski,M.-C., Darlix,J.-L. and Roques,B.P. (1991) *Biochem. Biophys. Res. Comm.*, **180**, 1010–1018.
25 Jentoft,J.E., Smith,L.M., Fu,X., Johnson,M. and Leis,J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 7094–7098.
26 Karpel,R.L., Henderson,L.E. and Oroszlan,S. (1987) *J. Biol. Chem.*, **262**, 4961–4967.
27 Tsuchihashi,Z. and Brown,P.O. (1994) *J. Virol.*, **68**, 5863–5870.
28 Bakalkin,G., Yakovleva,T., Selivanova,G., Magnusson,K.P., Szekely,L., Kiseleva,E., Klein,G., Terenius,L. and Wiman,K.G. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 413–417.
29 Oberosler,P., Hloch,P., Ramsperger,U. and Stahl,H. (1993) *EMBO J.*, **12**, 2389–2396.
30 Boehmer,P.E. and Lehman,I.R. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 8444–8448.
31 Alberts,B.M. and Frey,L. (1970) *Nature (London)*, **227**, 1313–1318.
32 Zijderveld,D.C. and van der Vliet,P.C. (1994) *J. Virol.*, **68**, 1158–164.
33 Khan,R. and Giedroc,D.P. (1992) *J. Biol. Chem.*, **267**, 6689–6695.
34 Echols,H. (1990) *J. Biol. Chem.* **265**, 14697–14700.
35 Coffin,J.M. (1979) *J. Gen. Virology*, **42**, 1–26.
36 Sykora,K.W. and Moelling,K. (1981) *J. Gen. Virol.*, **55**, 379–391.
37 Kumar,A. and Wilson,S.M. (1990) *Biochemistry*, **29**, 10717–10722.