*Nucleic Acids Research, 2006, Vol. 34, No. 2* ***593–605***
*doi:10.1093/nar/gkj458*

# Nucleic acid binding and chaperone properties of HIV-1 Gag and nucleocapsid proteins

**Margareta Cruceanu[1], Maria A. Urbaneja[2], Catherine V. Hixson[2], Donald G. Johnson[2], Siddhartha A. Datta[3], Matthew J. Fivash[4], Andrew G. Stephen[5], Robert J. Fisher[5], Robert J. Gorelick[2], Jose R. Casas-Finet[6], Alan Rein[3], Ioulia Rouzina[7] and Mark C. Williams[1,8,*]**

[1]Department of Physics, Northeastern University, 111 Dana Research Center, 110 Forsyth Street, Boston, MA 02115, USA, [2]AIDS Vaccine Program, SAIC-Frederick, Inc., NCI at Frederick, Frederick, MD 21702, USA, [3]HIV Drug Resistance Program, NCI-Frederick, Frederick, MD 21702-1201, USA, [4]Data Management Services, Inc., NCI-Frederick, Frederick, MD 2170, USA, [5]Protein Chemistry Laboratory, SAIC Frederick, Inc., NCI at Frederick, Frederick, MD 2170, USA, [6]MedImmune, Inc., 35 W. Watkins Mill Road, Gaithersburg, MD 20878, USA, [7]Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, 6-155 Jackson Hall, 321 Church Street SE, Minneapolis, MN 55455, USA and [8]Center for Interdisciplinary Research on Complex Systems, Northeastern University, 111 Dana Research Center, 110 Forsyth Street, Boston, MA 02115, USA

Received December 5, 2005; Revised and Accepted January 5, 2006

## ABSTRACT

**The Gag polyprotein of HIV-1 is essential for retroviral replication and packaging. The nucleocapsid (NC) protein is the primary region for the interaction of Gag with nucleic acids. In this study, we examine the interactions of Gag and its NC cleavage products (NCp15, NCp9 and NCp7) with nucleic acids using solution and single molecule experiments. The NC cleavage products bound DNA with comparable affinity and strongly destabilized the DNA duplex. In contrast, the binding constant of Gag to DNA was found to be ~10-fold higher than that of the NC proteins, and its destabilizing effect on dsDNA was negligible. These findings are consistent with the primary function of Gag as a nucleic acid binding and packaging protein and the primary function of the NC proteins as nucleic acid chaperones. Also, our results suggest that NCp7's capability for fast sequence-nonspecific nucleic acid duplex destabilization, as well as its ability to facilitate nucleic acid strand annealing by inducing electrostatic attraction between strands, likely optimize the fully processed NC protein to facilitate complex nucleic acid secondary structure rearrangements. In contrast, Gag's stronger DNA binding and aggregation capabilities likely make it an effective chaperone for processes that do not require significant duplex destabilization.**

## INTRODUCTION

The human immunodeficiency virus type 1 (HIV-1) contains two copies of single-stranded RNA (ssRNA), which are transcribed into double-stranded DNA (dsDNA) by a complex reverse transcription process. The linear dsDNA that results from this process is integrated into the DNA of the host cell, and from there it will be transcribed and translated using the gene expression machinery of the host cell. The main structural proteins of HIV-1 are expressed as a single polyprotein Gag (Figure 1A). Gag alone is sufficient for the formation of virus-like particles (VLP) in a mammalian cell (1). In fact, recombinant Gag, in the presence of nucleic acid and an additional small molecule cofactor, is sufficient for the assembly of VLP *in vitro* (2–4) which, except for the absence of a surrounding membrane, are morphologically very similar to the immature viruses formed *in vivo* (3,5). However, this self-assembly does not occur in the absence of nucleic acid (2,4,6,7). Surprisingly, the self-assembly of structurally normal VLPs has been shown to be supported by the presence of virtually any RNA or DNA molecules (2,4,6–8). However, the HIV-1 Gag protein is also capable of preferentially packaging viral RNA from a large pool of cellular nucleic acid

*To whom correspondence should be addressed. Tel: 1 617 373 7323; Fax: 1 617 373 2943; Email: mark@neu.edu
Correspondence may also be addressed to Ioulia Rouzina. Tel: 1 612 624 7468; Fax: 1 612 624 5121; Email: rouzina@cbs.umn.edu

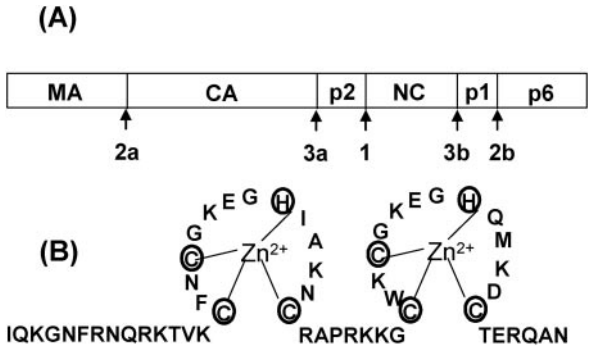


**Figure 1.** (**A**) Schematic representation of the cleavage steps that release the three forms of NC protein from Gag during viral maturation. Primary, secondary and tertiary cleavage sites are indicated by the numbered arrows, respectively. (**B**) Amino acid sequence of NCp7 (1–55) including its two-CCHC-zinc finger structures. NCp7 has 15 positive residues [5 arginines (R) and 10 lysines (K)] and two aromatic residues (Phe16 and Trp37). In addition, NCp9 (1–71) and NCp15 (1–123) have additional domains in the C-terminus which contain several aromatic residues, whose fluorescence, however, is not quenched upon binding of these proteins to nucleic acid.

molecules (9). This recognition appears to rely critically on the interaction of RNA with the two zinc fingers of the nucleocapsid (NC) domain and with its cationic residues (9–13), as shown in Figure 1B.

Translation of the Gag polyprotein from the mRNA derived from proviral DNA entails ribosomal frameshifting after the NC coding sequence in ~5% of cases (14,15). This frameshifting produces another polyprotein, Gag-Pol, which in addition to the Gag proteins also contains the viral enzymes protease (PR), reverse transcriptase (RT) and integrase (IN) (15). The self-processing of accumulating Gag-Pol, after virus assembly, results in growing levels of PR that begins the sequential cleavage of Gag (15). Extensive experiments have shown that the primary cleavage site of Gag by PR resides between the CA and NC domains after the p2 domain (site 1 in Figure 1A) (16,17). Two secondary cleavage sites, characterized by ~9-fold slower cleavage rates, reside between the MA and CA domains (site 2a) and between the p1 and p6 domains (site 2b) of the C-terminal part of Gag. Even slower cleavage events (~400-fold slower than the primary cleavage) occur at two additional sites, as the p2 and p1 peptides are cleaved from the C-terminal domains of CA and NC, respectively (sites 3a and 3b). In the fully mature, infectious virus, all proteins are present predominantly in their completely processed form (18). The three main products of the Gag processing by PR that contain the NC domain of Gag were termed according to their molecular weight and in the order of their appearance as follows: NCp15 (123 amino acids), NCp9 (71 amino acids) and NCp7 (55 amino acids) proteins (17,19).

All of the Gag cleavage events appear to be related to the changes in the structure and function of the virus during the course of its life cycle. Thus, the cleavage of NCp15, p2 and MA from Gag results in the formation of the cone-shaped capsid shell composed of CA, which is much smaller than the spherical membrane-bound MA shell (5,20,21). Within this cone-shaped mature core, there is still a dense phase of NC bound to the viral RNA dimer, a complex of the two RNA molecules within the virus (20,22,23). The RNA undergoes a process of further maturation, in which its secondary structure is altered. In a later stage of the retroviral life cycle, this RNA is reverse transcribed into proviral cDNA. The latter process involves several strand transfer steps. These events require major rearrangement of the nucleic acid secondary structure. While this restructuring results in a lower free energy of the nucleic acid molecules, and is therefore spontaneous, it appears to be extremely slow. NCp7 has been shown to significantly accelerate these reactions. The capability to facilitate the rearrangements required for these reactions is referred to as nucleic acid chaperone activity. However, it is not known to what extent the complete processing of Gag to its final form, NCp7 (Figure 1B) is required for the enhancement of specific rearrangements of nucleic acid secondary structure. For example, a recent study showed that while both NCp7 and Gag are effective at annealing $tRNA_3^{Lys}$ to the primer-binding site on the HIV-1 genome, Gag is more effective than NC (24). However, only processed NC has been shown to be capable of facilitating minus-strand transfer (25,26).

In order to understand the effects of the processing of NC on its nucleic acid chaperone and binding activities, we will examine the interaction of DNA with Gag, NCp7, NCp9 and NCp15 using a single DNA molecule stretching technique with optical tweezers as well as two standard solution techniques: tryptophan fluorescence and fluorescence anisotropy (FA) binding assays. In this combined approach, we are able to extract both traditional and novel information on DNA–protein interactions. First, we are able to determine the binding constants to the nucleic acid of the stoichiometrically bound proteins by DNA stretching, obtaining results that agree with both fluorescence studies. In addition, we are able to obtain information on the effect of each protein on force-induced DNA melting, i.e. strand separation due to an external pulling force. In turn, this enables us to determine the effect of each protein on DNA duplex stability and on the kinetics of the strand separation and annealing. While the nucleic acid binding properties of a protein can be compared with previous solution binding data, as is done here, the information on the effect of NC on DNA melting is unique and important for determining the nucleic acid chaperone properties of various forms of NC. Despite our use of long polymeric dsDNA, it is likely that our results are relevant for the binding and duplex destabilizing capabilities of NC to RNA as well. To demonstrate the relevance of our DNA stretching results to RNA binding, we compare binding of both Gag and NCp7 with similar short DNA and RNA oligonucleotides. The major conclusion of this study is that the nucleic acid chaperone activity improves greatly as the Gag precursor is progressively processed to create NCp15, NCp9 and NCp7.

## MATERIALS AND METHODS

### Protein preparation

NCp7, NCp9 and NCp15 from HIV-1 were expressed and purified as described (25,27,28). Briefly, HPLC fractions containing the expressed NC sequences (identified by MALDI-mass spectrometry and N-terminal sequencing) were collected, pooled and lyophilized. NC proteins were resuspended in the presence of 1 equivalent of $Zn^{2+}$ per finger,

and their concentration was calculated based on amino acid analysis measurements. Solutions were aliquoted into sterile polypropylene vials, relyophilized and stored at −70°C until used. The molecular masses of the proteins (as determined by MALDI-MS) were consistent with the following sequences: NCp7 = **IQKGNFRNQRKTVKCFNCGKEGHIAKNCRAPRKKGCWKCGKEGHQMKDCTERQAN**; NCp9 = NCp7 + $p1^{Gag}$ = **NCp7 + FLGKIWPSHKGRPGNF**; NCp15 = NCp7 + $p1^{Gag}$ + $p6^{Gag}$ = **NCp9 + LQSRPEPTAPPEESFRFGEETTTPSQKQEPIDKELYPLASLRSLFGSDPSSQ**.

Gag protein (lacking myristic acid at its N-terminus and the p6 domain at its C-terminus, referred to previously as Gag–Δp6) was purified from bacteria as described (2), except that EDTA was eliminated from the purification procedure so that zinc ions would be retained in the zinc fingers of the NC domain. This change required modification of several steps in the procedure, as the protein with zinc bound bacterial nucleic acids considerably more tightly than that lacking zinc. Gag concentrations were measured from their absorbance in 6 M guanidine hydrochloride, pH 6.5.

### Fluorescence binding studies

Poly(dT) was obtained in lyophilized form from Sigma (St Louis, MO). Poly(εA) was purchased from Pharmacia. LTR DNA oligonucleotides were purchased from Invitrogen with the following sequences: U5 sense, 5′-TAA CTA GAG ATC CCT CAG ACC CTT TTA GTC AGT GTG GAA AAT CTC TAG CAG T-3′; U5 antisense, 5′-ACT GCT AGA GAT TTT CCA CAC TGA CTA AAA GGG TCT GAG GGA TCT CTA GTT A-3′; U3 sense, 5′-ACT GGA AGG GCT AAT TCA CTC CCA AAG AAG ACA AGA TAT CCT TGA TCT GTG G-3′; U3 antisense, 5′-CCA CAG ATC AAG GAT ATC TTG TCT TCT TTG GGA GTG AAT TAG CCC TTC CAG T-3′.

The oligonucleotides were aliquoted in five OD vials by the manufacturer and lyophilized. Once dissolved, the concentration of each solution was checked by UV absorbance. To anneal the oligonucleotides in solution, equimolar amounts of the complementary strands were mixed in Eppendorf tubes, heated to near boiling and allowed to cool slowly. A spectrophotometric temperature ramp showed >20% hyperchromism for the annealed duplex. Sodium monophosphate and biphosphate, magnesium and sodium chloride were purchased from Aldrich Chemical Inc. (Milwaukee, WI).

Fluorescence titrations were carried out with a SPEX Fluoromax-2 spectrofluorimeter. NC Trp emission was monitored at 350 nm (5 nm bandwidth) with 280 nm excitation (1 nm bandwidth) in a thermostated capped quartz cell (0.2 × 1.0 cm; Hellma Cells, Inc., Jamaica, NY); stepwise additions of nucleic acid to 0.5–1.0 μM NC solutions, in 10 mM sodium phosphate, pH 7.0 were performed. Readings were corrected for dilution and inner filter effects. NC protein–nucleic acid complex formation was monitored by reduction in the initial fluorescence of the protein [or ethenoA fluorescence enhancement, in the case of poly(εA)].

NC fractional saturation was inferred from the ratio of observed quenching to maximal quenching (at saturation), $\Delta F_{lim}$, for each of the nucleic acid additions. Limiting quenching was calculated using the method of Kelly *et al*. (29). The plateau's intersection with the initial slope indicates the occluded binding site size (*n*) for a titration carried out under near-stoichiometric conditions such as that shown in Figure 2A. Polynucleotide binding cooperativity was analyzed using the model of McGhee and von Hippel (30).

Salt-back titrations were carried out by addition of a 5.0 M NaCl solution to a preformed NC–nucleic acid complex (Figure 2B). Double-logarithmic plots of log *K* versus log [$Na^+$] (Figure 2C) were used to derive binding constants under various ionic strength conditions (31). The binding free energy can thus be resolved into two components: an

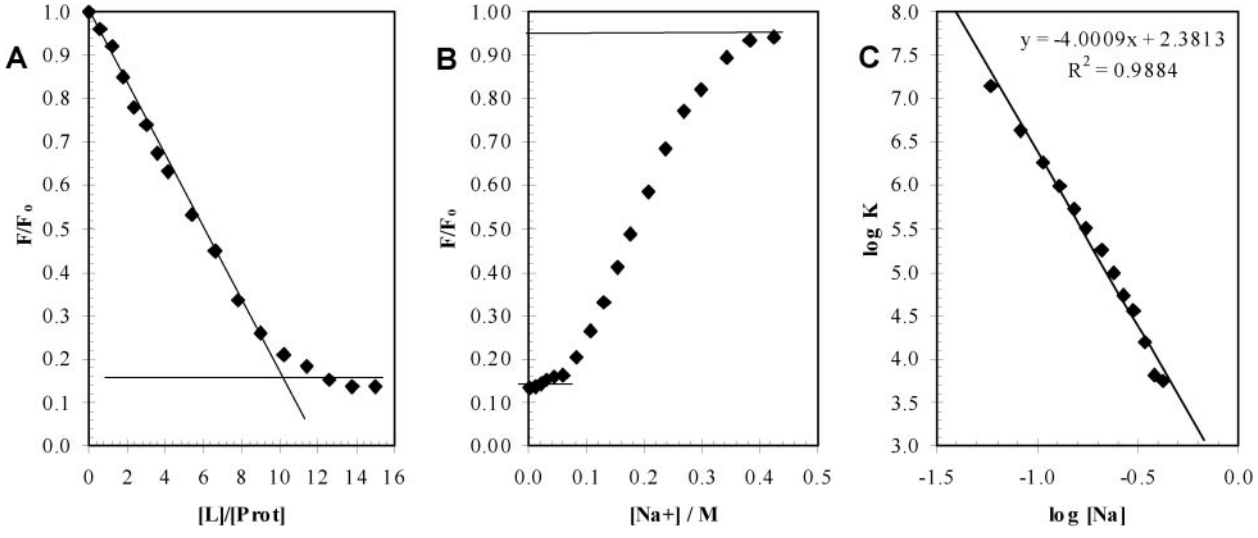


**Figure 2.** (**A**) Normalized fluorescence versus concentration ratio of nucleic acid to protein. Poly(dT) was added to NCp15 (0.5 μM in 1 mM sodium phosphate, pH 7.0, and 25°C). The plateau's intersection with the initial slope indicates the stoichiometry (*n*, occluded binding site size in the case of a polynucleotide) for a titration carried out under near-stoichiometric conditions such as that shown. (**B**) Fluorescence recovery ($F/F_0$) after the addition of 5 M NaCl aliquots to a pre-formed NCp15–poly(dT) complex. (**C**) Double-logarithmic plot for the association of NCp15 to poly(dT). The slope indicates the number of $Na^+$ ions displaced; the *y*-intercept is log $K^0$ at 1 M $Na^+$, when only non-ionic interactions contribute to binding free energy.

electrostatic contribution calculated from the slope of the regression line in the double-logarithmic plot, and a hydrophobic (or non-electrostatic) contribution from the *y*-intercept (32). In the absence of net anion uptake or release during binding, dividing the slope by the number of ions bound (in the thermodynamic sense) per phosphate group in the double-stranded or single-stranded nucleic acid lattice, yields the number of electrostatic interactions incurred by the protein. The *y*-intercept represents the logarithm of the binding affinity at 1 M $Na^+$ (log $K^0$), the standard state where electrostatic interactions are considered null. Low salt buffer was used in the initial titrations to maximize nucleic acid binding to NC proteins, permitting also a broader range of salt concentrations to be explored in the subsequent salt-back titration.

### Fluorescence anisotropy binding studies

FA is based on the observation that when a fluorescent molecule is excited with plane-polarized light, it emits polarized fluorescent light into a fixed plane if the molecules are stationary between excitation and emission (33,34). The FA is determined by measuring the ratio of the emission intensities from the parallel and perpendicular planes. Because the FA of a molecule is proportional to its molecular volume, this ratio is used to study molecular interactions.

$d(TG)_5$ labeled with fluorescein at the 3′ end was synthesized by Leo Lee, SAIC-Frederick and purified by HPLC. $(UG)_5$ labeled with fluorescein at the 3′ end was from Dharmacon (Chicago, IL). To determine the dissociation constants of Gag and NCp7 to these oligonucleotides, 10 nM stock solutions of either $d(TG)_5$ or $(UG)_5$ were prepared in HBS buffer (10 mM HEPES, 150 mM NaCl, pH 7.5) with 100 μM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), 5 mM β-mercaptoethanol and 1 μM $ZnCl_2$ and aliquoted into 96-well Costar polypropylene plates (Corning, NY). NC and Gag proteins were serially diluted into both oligonucleotides. An aliquot of 40 μl was removed and transferred into 384-well Costar polypropylene plates (Corning, NY). Samples were excited at 485 nm and the emission intensities at 535 nm from the parallel and perpendicular planes were measured using a Tecan Ultra plate reader (Durham, NC). Reactions were performed as three independent experiments. Curves were fit assuming a single binding site per oligonucleotide, using the relation (34):

$$An(C) = \frac{((R\cdot An_{\text{bound}} - An_{\text{free}})\cdot(\{(K_d + C + Lt) - [(K_d + C + Lt)^2 - 4\cdot Lt\cdot C]^{1/2}\}/2Lt) + An_{\text{free}})}{((R-1)\cdot(\{K_d + C + Lt) - [(K_d + C + Lt)^2 - 4Lt\cdot C]^{1/2}\}/(2Lt)) + 1)} \quad \mathbf{1}$$

where $An(C)$ is the measured anisotropy as a function of concentration, $Lt$ is the oligonucleotide concentration, $An_{\text{free}}$ is the anisotropy of unbound oligonucleotide, $An_{\text{bound}}$ is the anisotropy of the completely bound oligonucleotide solution, $R$ is the ratio of the fluorescence intensity of saturated bound oligonucleotide relative to free oligonucleotide, $C$ is the protein concentration and $K_d$ is the measured dissociation constant resulting from the fit. Because we do not actually know that there is only a single binding site per oligonucleotide, below we refer to the $K_d$ measurements resulting from this fit as 'apparent $K_d$'.

### Force-induced melting

The optical tweezers instrument and data acquisition were described previously (35,36). Biotinylated DNA was obtained as previously described (37). The DNA was labeled at the 3′ ends only such that it would be free to rotate when stretched. All proteins were aliquoted in volumes of 10 μl and 10 or 25 μM protein concentration in specific NC-storage buffer (20 mM HEPES, 5 mM β-mercaptoethanol, 0.1 mM TCEP, pH 7.5), and stored at −80°C before use. NC proteins were reconstituted with two molar equivalents of zinc. The tethering buffer used in this study for all force-induced melting experiments (FIM buffer) contained 10 mM HEPES with 45 mM NaCl and 5 mM NaOH, pH 7.5. Lambda DNA, dNTPs (dATP, dTTP, dGTP), enzymes and buffers were purchased from Promega, the biotin-11-dCTP from Enzo, HEPES, NaCl and NaOH from Sigma, and the streptavidin-coated beads from Bang Laboratories. After capturing a single DNA molecule in the tethering buffer, the molecule was stretched to verify that the correct force-extension curve was obtained in the absence of bound protein. To measure the effect of the proteins on this transition, a buffer solution containing a fixed concentration of protein dissolved in FIM buffer was added to the experimental cell until the buffer surrounding the captured DNA molecule was completely exchanged.

## RESULTS

Below we will present our results in two main sections. In the first, we use a single DNA molecule stretching technique and in the second we use a tryptophan fluorescence assay. These experiments are designed to elucidate the specific features of the interaction with DNA of the HIV-1 NC protein while in the context of Gag as well as in its three differently processed forms. The overall goal is to understand how the sequential proteolysis of HIV-1 NC affects its nucleic acid interaction properties, and to determine whether this can be correlated with the change of NC's function in various stages of the viral replication cycle. In a third section, we compare the binding of NC and Gag with a very short oligoribonucleotide and a very short oligodeoxynucleotide to determine whether the 2′ OH (or the methyl group in thymidine) plays a significant role in the interactions of these proteins with nucleic acids.

### Single DNA molecule stretching studies of NC–DNA and Gag–DNA interactions

In this work we apply a single DNA molecule stretching technique to obtain information on DNA interactions with each of the HIV-1 NC and Gag proteins. Some information about the NCp7–DNA interaction obtained by this method was already discussed in our previous work (35,36). In this report we perform a much broader analysis of the HIV-1 NC–DNA interaction as a function of protein concentration, which allows us to determine the binding constant of each protein to DNA. In addition, we analyze the equilibrium and non-equilibrium features of our stretching curves, which allow us to determine both equilibrium and kinetic properties of protein–DNA interactions.

An equilibrium DNA stretching curve, $F(x)$, is extremely sensitive to the presence of DNA binding proteins, such as NC, that bind stoichiometrically along the whole length of the DNA molecule. It, therefore, can be used to monitor this binding. To do this, we obtain the stretching curves of individual DNA molecules undergoing the force-induced melting transition. At large enough extensions the DNA under tension is a combination of melted and double helical regions, with the relative proportion of both forms varying with the molecular extension. Because the HIV-1 NC proteins are known to bind dsDNA and single-stranded DNA (ssDNA) with comparable affinities (Table 2), we will not try to distinguish between these modes of binding, but instead use our $F(x)$ curves in order to obtain an estimate of the average composite binding constant to both forms of DNA. While this method cannot separately determine dsDNA and ssDNA binding, it allows us to quantify the interaction of NC with polymeric natural DNA. Such DNA more strongly resembles the actual substrate for NC binding *in vivo* relative to other artificial constructs such as oligonucleotides and homopolymers. This method was used successfully in a previous work to obtain binding constants for LINE-1 ORF1p and several mutants, and the results obtained were in agreement with bulk binding experiments (38).

In order to estimate the equilibrium binding constant of the protein to DNA, we can follow the variation of the force-extension curve $F(x)$ with the protein concentration $c$. The most distinct feature of the changes induced by the presence of the NC proteins in DNA stretching and relaxation behavior is the significant increase in the slope of the force-induced melting transition upon addition of the protein. This increase in transition width is likely due to NC's sequence-specific binding and subsequent reduction in DNA melting cooperativity, as well as intercalation of NC into dsDNA upon stretching. All of these effects contribute to NC's nucleic acid chaperone capabilities. Here, we use this growing transition width as a quantitative parameter to monitor the protein binding to ssDNA. We can define the transition width $\Delta F$ as a force interval over which the DNA molecule switches from an entirely ds state to an entirely ss state. To estimate this transition width, we fit a tangent line to the transition midpoint (determined by linear regression analysis of several stretching curves), as shown in Figure 3. This midpoint corresponds to half the extension change between dsDNA only (solid line, closed circle), and ssDNA only (solid line, closed triangle) in the presence of protein. The slope of this line is the change in force from the beginning of the transition to the end of the transition, divided by the change in DNA length during the transition. Thus, the force transition width $\Delta F$ is the difference in forces at which this tangent line intersects tangent lines to the dsDNA and ssDNA extension curves (Figure 3). The values obtained in this way are plotted as a function of protein concentration $c$ for all four proteins studied (Figure 4). In all cases the transition width first increases linearly with protein concentration, and saturates at $\Delta F_{sat}$. We see that the plots for all three forms of NC protein used in this study are similar within the experimental errors and very different from the plot for Gag. The slope of the linear region of this plot can be used to estimate the binding constant. Indeed, at sub-saturating protein concentrations, the fraction $\Theta$ of DNA binding sites bound by protein is very small, $\Theta \ll 1$, and increases linearly

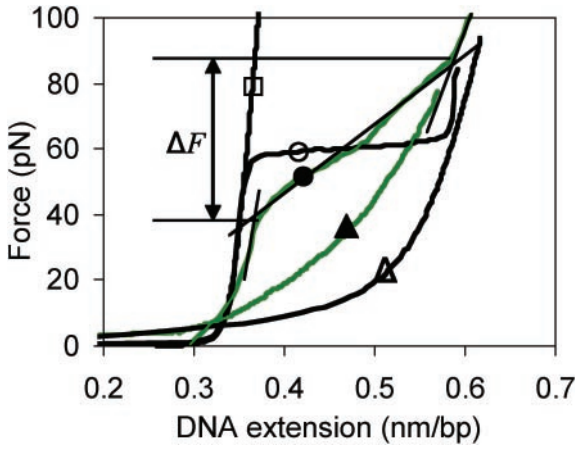


**Figure 3.** Stretching curves for λ-DNA in the absence of protein (open circle), in the presence of 20 nM NCp7 (wild type) protein (closed circle), and for ssDNA with NC bound on it (closed triangle), all in FIM buffer (see Materials and Methods) and room temperature. For ssDNA with NC bound, data are from ref. (35), and was taken with 10 nM NCp7 in 25 mM NaCl, 10 mM HEPES, pH 7.5, room temperature. Stretching curves for dsDNA (open square) and ssDNA (open triangle), fit to standard polymer models, in the absence of protein (90) are also shown. The method used to determine $\Delta F$, the force transition width, is shown schematically.

with the concentration of the protein $c$:

$$\Theta = K \cdot c, \qquad \mathbf{2}$$

where $K$ is the apparent binding constant of protein to DNA. Note that the free protein concentration does not depend on the extent of binding because the protein is in great excess over the DNA in this experiment. In addition, NCp7 is known to exist as a monomer in solution even at high concentrations (39), and while Gag can form dimers in solution, at the concentrations used here the number of dimers should be negligible (S. A. Datta and A. Rein, unpublished data). Assuming that the change in the transition width, $\delta F = \Delta F - \Delta F_0$, is proportional to the fraction of sites bound by protein, $\Theta$, and with the saturated value of this parameter, $\delta F_{sat}$, we can also write:

$$\delta F(c) = \Theta \cdot \delta F_{sat} \qquad \mathbf{3}$$

Here, $\Delta F_0$ is the transition width in the absence of protein. Combining Equations 2 and 3 we can estimate the protein–DNA equilibrium binding constant as:

$$K = \frac{1}{\delta F_{sat}} \cdot \frac{\delta F(c)}{c} = \frac{1}{\delta F_{sat}} \cdot \frac{d(\delta F(c))}{dc}\Big|_{c \to 0} \qquad \mathbf{4}$$

The fact that these stretching curves are not completely reversible does not preclude the measurement of equilibrium binding isotherms. The stretching curves used for these measurements did not depend on pulling rate (at the low pulling rates used here) and the measurements therefore are indistinguishable, within the reported experimental error, from equilibrium measurements (40). Note that irreversible thermal melting curves (which do not depend on heating rate) are routinely used to obtain equilibrium thermodynamic measurements (41).

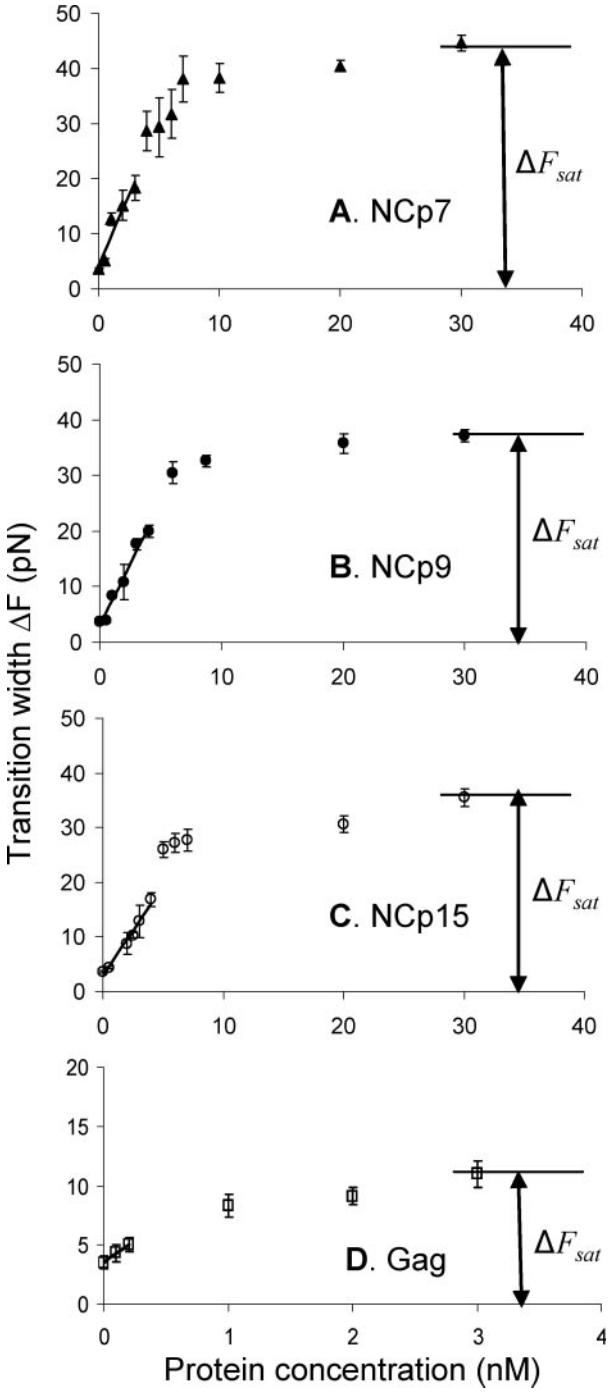


**Figure 4.** Dependence of the transition width $\Delta F$ on the protein concentration. Data are for: (**A**) NCp7, (**B**) NCp9, (**C**) NCp15 and (**D**) Gag. Each data point for the NC proteins is given as the mean ± standard error for seven experiments, and with fewer experiments for Gag. Data are taken in FIM buffer (see Materials and Methods), and room temperature. The solid lines are linear fits to the data in the low concentration limit, as follows: NCp7: $\delta F = (5.02 \pm 0.97)\cdot c$; NCp9: $\delta F = (4.69 \pm 0.58)\cdot c$; NCp15: $\delta F = (3.29 \pm 0.22)\cdot c$; Gag: $\delta F = (7.54 \pm 0.93)\cdot c$. Here, $\delta F$ is the difference between $\Delta F$ in the presence of protein and $\Delta F$ in the absence of protein.

The parameters obtained when fitting the experimental data for all proteins studied (Figure 4) are shown in Table 1. We see that while the binding constants for the three processed forms of NC protein are indistinguishable, $K \sim 1 \times 10^8$ M$^{-1}$,

**Table 1.** Results of force-induced melting experiments in the presence of HIV-1 NCp7, NCp9, NCp15 and Gag

| Protein | $\Delta F_{sat}$[a,b] (pN) | $d(\delta F)/dc$[a,c,d] (pN/nM) | $K$[e] ($10^8$ M$^{-1}$) | $\Delta G_{total}$[a,f] ($k_BT$/bp) | $\Delta G_{total}$[a,f] (kcal/mol bp) |
|---|---|---|---|---|---|
| None | 3.66 ± 0.16 | — | — | 2.31 ± 0.05 | 1.36 ± 0.03 |
| NCp7 | 48.4 ± 1.9 | 5.02 ± 0.97 | 1.22 ± 0.25 | 0.65 ± 0.02 | 0.38 ± 0.01 |
| NCp9 | 40.9 ± 1.1 | 4.69 ± 0.58 | 1.40 ± 0.19 | 0.58 ± 0.02 | 0.34 ± 0.01 |
| NCp15 | 39.1 ± 1.6 | 3.29 ± 0.22 | 1.04 ± 0.09 | 0.57 ± 0.02 | 0.33 ± 0.01 |
| Gag | 11.5 ± 2.0 | 7.54 ± 0.93 | 9.43 ± 2.6 | 1.86 ± 0.07 | 1.08 ± 0.04 |

[a]Measurements were performed in FIM buffer (see Materials and Methods).
[b]$\Delta F_{sat}$ is the measured transition width at 30 nM HIV-1 NC or 3 nM Gag concentration (saturation) as shown in Figure 4.
[c]$\delta F$ is defined as the difference between $\Delta F$ in the presence of protein and $\Delta F$ in the absence of protein.
[d]$d(\delta F)/dc$ is the slope of the linear fit to experimental data in the low binding fraction limit (solid lines in Figure 4).
[e]$K$ is calculated using Equation 4.
[f]The melting transition free energy $\Delta G_{total}$ is calculated as the area between the experimental stretching curve for dsDNA without (open circle)/with (closed circle) protein at saturated transition width (30 nM NC or 3 nM Gag) and the experimental stretching curve for the ssDNA without (open triangle)/with (closed triangle) protein bound on it (Figure 3).
Data are reported as mean ± standard error, $n \geqslant 3$. Averages were obtained for seven experiments. Errors for the slopes are calculated using the regression formula. The helix-coil transition free energy $\Delta G$ is reported in units of $k_BT$ or kcal/mol per base pair (bp), where $k_B$ is the Boltzmann constant and $T = 293$ K is room temperature ($1\ k_BT = 581$ cal/mol $= 4 \times 10^{-21}$ J).

the saturated value of the transition width increases slightly from NCp15 to NCp7. That is, $\Delta F_{sat} = 35$, 37 and 45 pN for NCp15, NCp9 and NCp7, respectively. On the other extreme, the Gag protein has higher binding affinity for ssDNA, $K \sim 9 \times 10^8$ M$^{-1}$, but the increase in the transition width is much lower, such that $\Delta F_{sat} \sim 11$ pN.

In addition to measuring the binding constant of each protein to DNA, our DNA stretching curves can be used to determine the dsDNA melting free energy, as discussed previously (42,43). In this report we will outline this approach and then use it for all four proteins at saturated binding in order to characterize the maximum effect of these proteins on DNA duplex stability. Briefly, the area of the mechanical cycle in the $(F - x)$ plane delimited by the experimental stretching curve and the extrapolated stretching curve for ssDNA, with or without protein bound, gives the melting free energy in each case. Note that if the extension is per base pair, then the free energy is also per base pair, averaged over the whole length of polymeric DNA. The mechanical energy required to melt an average base pair, in turn, equals the free energy of the duplex relative to the single-stranded state of a base pair in the absence of applied force. The latter value is a function of solution conditions, including the concentration of DNA-binding protein. This method allows us to measure directly the effect of the protein on the stability of the duplex. Importantly, this approach does not rely on any model of the protein–DNA interaction or on the details of the DNA melting transition. We note that when the stretching curves are not reversible, the work done by stretching may not represent the equilibrium melting free energy. If that is the case, our results will slightly overestimate the melting free energy. However, this effect should be very small due to the relatively high cooperativity of DNA melting under the conditions studied here as well as the observed rate-independence (at

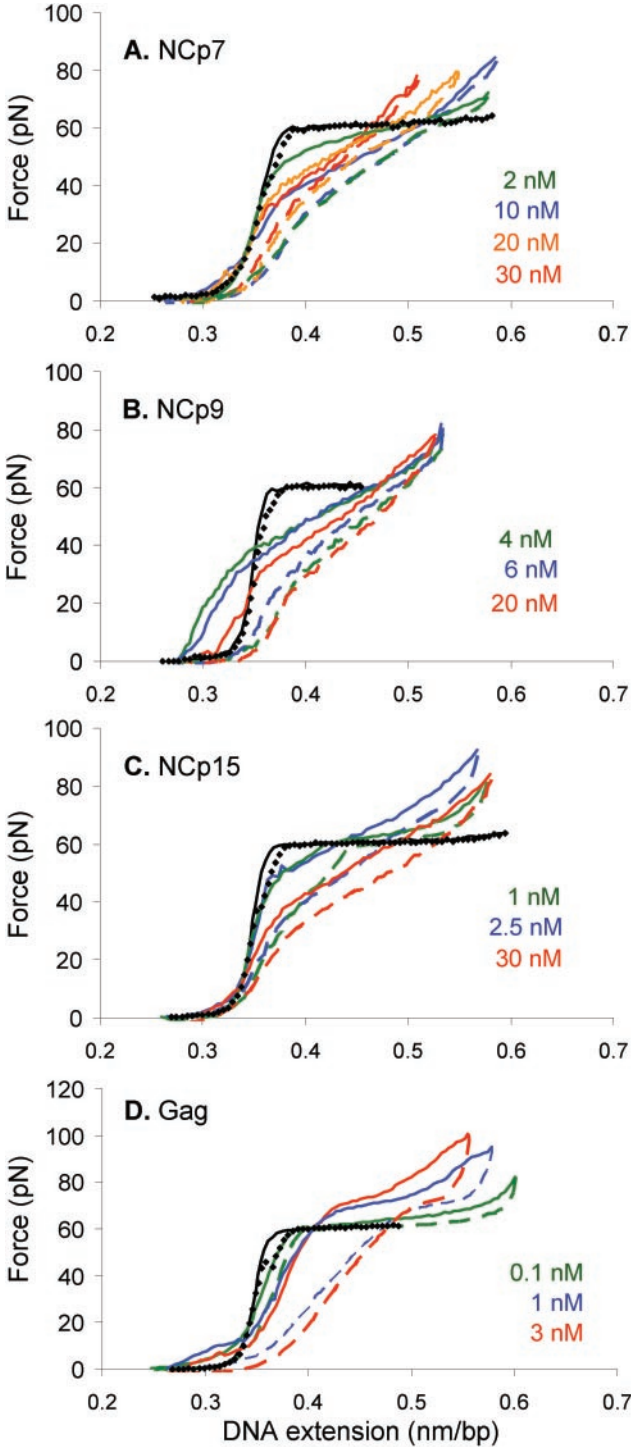


**Figure 5.** Stretching (solid lines) and relaxing (dashed lines) curves for λ-DNA in the presence of the NC proteins or Gag: (**A**) NCp7; (**B**) NCp9; (**C**) NCp15; (**D**) Gag. Protein concentrations are indicated in the figure. Data were collected in FIM buffer (see Materials and Methods) at room temperature. The stretching and relaxing curves in the absence of protein are also shown (black solid line and filled diamonds, respectively).

the low pulling rates used here) of the melting forces in these experiments.

Representative DNA stretching and relaxation curves for all four proteins are shown in Figure 5. The stretching portion of the curve is used to obtain the DNA melting free energy per base pair in the presence of each protein, calculated at saturated protein binding. The results of this calculation are presented in Table 1. The addition of saturating amounts of any of the three NC proteins (20 nM) results in a low DNA helix stability of 0.35 ± 0.03 kcal/mol bp. This value is about 1 kcal/mol bp smaller than in the absence of protein at the 50 mM ionic strength used. The significant duplex destabilization induced by NC is the result of the preferential interaction between its zinc fingers and the single-stranded nucleic acids. In addition, NC produces some duplex stabilization due to increased electrostatic screening of DNA phosphate charges. The latter effect depends on the ionic strength of the buffer used and results in a salt dependence of the total NC-induced duplex destabilization. Thus, at higher salt, the net NC-induced DNA destabilization can become slightly larger than 1 kcal/mol bp. The effect of Gag on duplex stability is less significant, i.e. the free energy of base pair melting is only 0.3 kcal/mol bp lower than in the absence of NC, due to the overall increase in the melting force and small change in the transition width (Figure 5). One of the possible reasons why Gag does not significantly destabilize the DNA duplex could be that its NC domain is not capable of interacting with the ssDNA bases via its zinc fingers as efficiently as in the processed NC. This possibility, however, is contradicted by the results of our direct measurement of NC and Gag binding to short ss oligonucleotides, discussed below, showing identical NC and Gag binding. The alternative explanation is that the binding of Gag molecules to dsDNA is always accompanied by strong interaction between the neighboring Gag molecules. The latter may produce a torsional constraint on the local unwinding of dsDNA, thus stabilizing it relative to the unwound ssDNA. This hypothesis is supported by our observation of the increasing force at the midpoint of the DNA force-induced melting transition upon Gag binding. Such a strong increase in the DNA melting force was previously observed when the stretched dsDNA was torsionally constrained (44). Thus, our results argue that the Gag polyprotein is inefficient as a duplex destabilizer, in contrast to all processed forms of NC. This finding suggests that Gag should be much less efficient at complex nucleic acid secondary structure rearrangements than the processed NC proteins.

Figure 5 also shows DNA relaxation curves for all four proteins, all of which show different amounts of hysteresis (i.e. disparity between stretch and relax curves). We have also studied the complete stretch–relax cycles of λ-DNA in the presence of many other retroviral NC proteins (M. Cruceanu, K. Stewart, R.J. Gorelick, K. Musier-Forsyth, I. Rouzina, M.C. Williams, manuscript in preparation), as well as of several mutants of HIV-1 NCp7 (M. Cruceanu, M.-N. Vo, R.J. Gorelick, K. Musier-Forsyth, I. Rouzina, M.C. Williams, manuscript in preparation), and the ORF1 protein from mouse LINE-1 retrotransposon (38). Among all DNA binding proteins studied using force-induced melting of DNA, including the proteins presented herein, HIV-1 NCp7 protein produces the smallest hysteresis of the DNA stretch–relax cycle. As shown in Figure 5, the hysteresis induced by the proteins studied here, at saturation, increased in the following order: HIV-1 NCp7, NCp15, NCp9 and Gag. However, while the hysteresis contains rich information on protein–DNA interactions, its interpretation is complex, and has never

been previously discussed. In Discussion, we present comments on the general information that can be obtained from studying the hysteresis of single DNA molecule stretching, followed by several conclusions regarding the proteins studied in this work. We will show that, although the information from studying the non-equilibrium features of the DNA stretch–relax cycle is not quantitative, it can serve as an important basis for understanding the kinetics of protein–nucleic acid interactions.

## Fluorimetric titrations with HIV-1 NC proteins

The nucleic acid binding properties of HIV-1 NC in its three differently processed forms were investigated with the long single-stranded homopolymeric DNA and RNA molecules, poly(dT) and poly(εA), respectively (45). Both polynucleotides were bound by the proteins nearly stoichiometrically under the conditions of the reverse titration experiment (see Materials and Methods). Binding to poly(dT) resulted in a large reduction (86%) in the fluorescence emission of the sole Trp residue (Trp37) in NCp7, as reported previously (45,46). The presence of the additional Trp residue (Trp61) in NCp9 and NCp15 resulted in a decrease in the limiting quenching at saturation for these NC sequences (Table 2), in agreement with earlier reports that Trp61 is not directly involved in the interaction of NCp9 with nucleic acids (46). Indeed, the binding affinity and thermodynamic parameters derived from salt-back titrations were nearly identical for NCp7 and NCp9. The occluded binding site size increased from NCp7 to NCp9 to NCp15, reflecting the larger size of the binding protein. NCp15 bound poly(dT) with an affinity an order of magnitude larger than NCp7 or NCp9. However, this effect was not observed for poly(ϵA) titrations. The binding free energy in the standard state (1 M NaCl) was comparable for all three proteins, indicating that the binding affinity of NCp15 would tend to converge with that of NCp7 and NCp9 at the higher ionic strength typical of physiological conditions.

dsDNA binding was also examined with long terminal repeat (LTR) sequences (Table 2). An increase in occluded binding site size was again seen in the association of NCp9 with duplex DNA, compared with that of NCp7. Owing to the scarcity of the material, NCp15 binding titrations were not carried out. Somewhat lower limiting quenching levels were obtained for NCp7 and NCp9, as compared with the related single-stranded sequences, in agreement with results reported for NCp7 using other duplex DNA molecules (47). Trp quenching was again highest for the single-Trp NCp7, pointing to a lack of involvement of the second Trp residue in NCp9 or NCp15, in agreement with their similar non-electrostatic contribution (log $K^0$) to the binding free energy, regardless of the presence of the C-terminal segment containing Trp61. Binding of all proteins to double-stranded LTRs had a larger electrostatic component than to ssDNA and was characterized by a smaller site size. NCp9 bound almost an order of magnitude tighter than NCp7 to LTR sequences. Overall, the bulk fluorescence titration results confirm the findings from the stretching experiments, which is that the equilibrium binding affinities of NCp7, NCp9 and NCp15 to nucleic acids are of similar magnitude.

**Table 2.** Nucleic acid binding properties of HIV-1 NC proteins[a]

| System | $n_{av}$ | $\Delta F_{lim}$ (%) | $K$ ($M^{-1}$) | dlog $K$/ dlog [$Na^+$] | log $K^0$ |
|---|---|---|---|---|---|
| NCp7 + poly(dT) | 6.4 | −86 | $9.3 \times 10^7$ | −2.54 | 2.89 |
| NCp9 + poly(dT) | 9.5 | −83 | $1.1 \times 10^8$ | −2.61 | 2.83 |
| NCp15 + poly(dT) | 10.3 | −70 | $1.4 \times 10^9$ | −3.35 | 2.45 |
| NCp7 + poly(εA) | 6.6 | +135 | $1.4 \times 10^9$ | −2.02 | 5.54 |
| NCp9 + poly(εA) | 10.2 | +210 | $3.8 \times 10^9$ | −2.06 | 5.11 |
| NCp15 + poly(εA) | 11.8 | +195 | $1.7 \times 10^9$ | −2.04 | 5.40 |
| NCp7 + 3′LTR-dsDNA | 2.7 bp | −85 | $8.3 \times 10^8$ | −3.26 | 2.40 |
| NCp9 + 3′LTR-dsDNA | 4.3 bp | −61 | $3.8 \times 10^9$ | −3.65 | 2.28 |
| NCp7 + 5′LTR-dsDNA | 2.9 bp | −86 | $3.4 \times 10^8$ | −3.21 | 2.49 |
| NCp9 + 5′LTR-dsDNA | 4.4 bp | −59 | $8.9 \times 10^9$ | −3.96 | 2.03 |
| NCp7 + 3′LTR/U5 | 6.6 | −92 | $4.3 \times 10^{7b}$ | −2.31 | 3.30 |
| NCp9 + 3′LTR/U5 | 8.3 | −81 | $6.2 \times 10^{7b}$ | −2.30 | 3.46 |
| NCp7 + 5′LTR/U3 | 6.8 | −93 | $2.2 \times 10^{8b}$ | −2.96 | 2.77 |
| NCp9 + 5′LTR/U3 | 9.5 | −86 | $3.9 \times 10^{8b}$ | −2.85 | 3.23 |

[a]Thermodynamic results from fluorimetric binding isotherms of NCp7, NCp9 and NCp15 with various polynucleotides were obtained by two methods: (i) addition of nucleic acid aliquots to protein (Trp quenching) and (ii) addition of protein aliquots to poly(εA), both in a solution of 10 mM sodium phosphate buffer, pH 7.0, at 25°C. Average errors were as follows: ±1 for $n_{av}$; ±1.6% for $\Delta F_{lim}$; ±28% for $K$; ±0.07 for log($K$)/log([$Na^+$]); and ±0.19 for log ($K^0$). $K^0$ is the association constant extrapolated to 1 M salt.
[b]These results were measured only at higher salt than that reported for the other measurements in the table. To allow for comparison of results, they have been extrapolated to 10 mM sodium phosphate buffer conditions.

## Fluorescence anisotropy binding studies with Gag and NCp7

Because the DNA stretching studies were done with DNA only, we wanted to test whether or not our conclusions concerning Gag and Gag cleavage products containing NC were applicable to Gag and NC binding to RNA. The interactions of NC with either RNA or DNA are both potentially biologically important in the context of NC-induced facilitation of reverse transcription. To this end, we performed FA binding studies of Gag and NCp7 binding to 10 nt DNA and RNA oligonucleotides, $d(TG)_5$ and $r(UG)_5$. The results of these measurements are presented in Table 3. The measured values of the binding constant for NC are in agreement with its value obtained in our stretching and tryptophan fluorescence quenching studies (see Tables 1 and 2). It seems likely that this agreement is in fact a result of mutual compensation of two effects. The FA binding studies were performed at higher (150 mM NaCl) salt, as compared with 50 mM NaCl in DNA stretching and 10 mM NaCl in tryptophan quenching experiments. Higher salt decreases the binding affinity of NC to nucleic acids. However, this weakening is compensated by the fact that TG and UG oligonucleotide sequences bind NC more strongly than random nucleic acid sequences (48,49). Perhaps the most important observation is that both NC and Gag bind similarly to RNA and DNA

**Table 3.** Binding constants of HIV-1 NCp7 and Gag proteins to 10 nt ssRNA and ssDNA oligomers measured by FA

| Oligonucleotide | $K$ ($M^{-1}$) $(UG)_5$ | $K$ ($M^{-1}$) $(TG)_5$ |
|---|---|---|
| NCp7 | $(8.1 \pm 1.2) \times 10^7$ | $(7.2 \pm 1.5) \times 10^7$ |
| Gag | $(5.2 \pm 2.2) \times 10^7$ | $(3.4 \pm 0.9) \times 10^7$ |

Binding measurements were performed as described in Materials and Methods.

oligomers. This result is consistent with previous observations of comparable NC binding to DNA and RNA molecules (49,50).

Another very important result is that the binding constants of NC and Gag to these short oligonucleotides are very similar, see Table 3. This result is in contrast to 10-fold stronger binding of Gag to long dsDNA, observed in our stretching study (Table 1). One possible explanation for this discrepancy is that Gag binds the longer dsDNA by some additional region besides NC, most likely by the matrix domain of Gag (MA). The alternative explanation is that the longer nucleic acid molecule is capable of binding several Gag molecules simultaneously, while the short oligonucleotides used (10 nt) can likely accommodate only one Gag molecule. In this case, even if the individual Gag–DNA and NC–DNA binding constants are equally strong, Gag–Gag interactions can contribute to the stronger apparent Gag binding observed in DNA stretching experiments. In Discussion below we explore this hypothesis and estimate the DNA binding cooperativity parameter of Gag, and the corresponding free energy of Gag–Gag interaction.

## DISCUSSION

In this work we used a single DNA molecule stretching technique, as well as a tryptophan fluorescence assay, in order to understand the specific features of the interaction with DNA of the HIV-1 NC protein while in the context of Gag as well as in its three differently processed forms. Our major goal is to see how the sequential proteolysis of HIV-1 NC affects its nucleic acid interaction properties, and whether or not this can be correlated with the change of NC's function in various stages of the viral life cycle. While the NC protein within the context of Gag appears to have mostly a nucleic acid binding and packaging function, supporting Gag self-assembly into VLP, the processed forms of NC appear to act mostly as nucleic acid chaperones. The latter finding means that NC is capable of greatly facilitating nucleic acid restructuring processes that lead to the more stable annealed state of these molecules. While Gag facilitates $tRNA_3^{Lys}$ annealing to the HIV-1 primer binding site (51–53), it has not been shown to facilitate more complex nucleic acid rearrangements, such as those involved in strand transfer. Until recently, it was not clear which properties of NC are responsible for its chaperone activity. However, recent studies of the kinetics of annealing (54), strand transfer (25,55–58) and recombination (59–62) have shown that the chaperone activity of NC has two major components: helix destabilization and strand aggregation (63).

HIV-1 NC is able to weakly destabilize all nucleic acid base pairs, allowing it to melt unstable nucleic acid intermediates, but preventing melting of the final, stably annealed structures. Thus, melting by NC of the short helical regions bordered by mismatches or bulges provides the single-stranded nucleation sites for the new helices. In addition, NC-induced duplex destabilization accelerates various processes of strand exchange (63). On the molecular level, the capability of HIV-1 NC to destabilize nucleic acid duplexes relies on its preferential binding to single-stranded nucleic acids *via* the stacking of the aromatic residues of its zinc fingers with unpaired bases (49,64–66). In our force-induced melting studies, this duplex destabilizing activity of NC is observed directly by measuring the changes in DNA base pair stability upon addition of NC. In the present work this duplex destabilizing activity of all three NC proteins studied appeared to be quite similar, with each base pair destabilized by processed NC proteins by ~1 kcal/mol bp. Since the average stability of a base pair in the absence of NC was ~1.4 kcal/mol in 50 mM NaCl, this is a significant but incomplete destabilization induced by NC, which is consistent with the chaperone function of NC. Similar destabilization produced by all three proteins seems reasonable, taking into account the fact that they all share the same DNA binding domains.

In contrast, according to our measurements, the DNA binding constant of Gag is about 10-fold higher than that of the other three NC proteins. If the NC domain constitutes the main binding domain of Gag to DNA, the enhancement of the protein–DNA association constant comes most likely from an additional Gag–Gag interaction upon binding of the protein to DNA (i.e. cooperative binding) (30), although we cannot rule out additional interactions from other parts of Gag. In the limit of low protein binding, $\Theta \ll 1$, the fraction of protein bound $\Theta$, is proportional to the protein concentration $c$, as in Equation 2:

$$\Theta = K^{\text{eff}} \cdot c, \qquad \mathbf{5}$$

and the effective binding constant includes the cooperativity parameter, $\omega$, defined below:

$$K^{\text{eff}} = K \cdot \omega \qquad \mathbf{6}$$

$$\omega = \exp\left[\frac{(m/2) \cdot \delta G_{\text{Gag-Gag}}}{k_B T}\right] \qquad \mathbf{7}$$

Here $\delta G_{\text{Gag-Gag}}$ is the free energy of Gag–Gag interaction per protein molecule, $m$ is the number of the nearest-neighbor Gag molecules when bound to DNA, the factor 1/2 takes into account that each interaction is shared by the two Gag molecules, and $k_B T$ is the thermal energy per statistical degree of freedom, which at room temperature is ~0.6 kcal/mol. The fact that the effective binding constant of Gag is ~10-fold higher than for either NC protein implies that $\omega \sim 10$. This estimate, of course, assumes that the intrinsic binding constant of Gag to DNA, $K$, is the same as the binding constant of NC to DNA. By making this assumption and using Equation 7, we can estimate that

$$m/2 \cdot \delta G_{\text{Gag-Gag}} = k_B T \ln \omega \approx 2.3 k_B T = 1.4\,\text{kcal/mol}. \qquad \mathbf{8}$$

The pair-wise attraction free energy $\delta G_{\text{Gag-Gag}}$ can be estimated from the net free energy of cooperative interactions per Gag molecule given by Equation 8, for the particular number of nearest neighbors, $m$. The $m$ range for Gag is likely from 3 to 4, with individual interacting subunits of capsid forming dimers of trivalent capsid molecules (67–69), each of which has 2–3 neighbors in addition to the dimer partner. In our experiments the number of interacting Gag neighbors bound to the single stretched DNA molecules could be as low as 2, given the fact that the DNA is stretched. We can therefore estimate the Gag–Gag interaction free energy to range between $\delta G_{\text{Gag-Gag}} = 0.7$ and 1.4 kcal/mol.

These results suggest that the individual CA–CA interactions of HIV-1 Gag are much weaker than any typical functionally important protein–protein interaction. Unfortunately, other direct measurements of such weak interactions are not yet available for HIV-1 Gag. However, the cooperative Gag–DNA binding constant that we infer from our results, $K_{Gag} = 10^9\ M^{-1}$ is in agreement with the typical values of the reciprocal critical concentration of capsid molecules required for the capsid self-assembly $K_{ap} = 10^6$–$10^9\ M^{-1}$ for several other viruses studied (70). Thus, cooperative binding of Gag to DNA is consistent with its enhanced capability to bind to and package nucleic acids. However, multiple Gag–Gag interactions are expected to lead to a slowing of this protein's motion on nucleic acid. This, in turn, should result in the interference of Gag with the fast nucleic acid duplex melting/annealing observed in the presence of NCp7, as reflected in the significant hysteresis observed in DNA stretching curves in the presence of Gag (Figure 5). Therefore, because of this slow unbinding (relative to NCp7) as well as the inability of Gag to destabilize the DNA duplex, the interactions of Gag with nucleic acids are not likely to be optimized for complex rearrangements of nucleic acid secondary structure such as those required for strand transfer.

The second HIV-1 NC activity that contributes to its chaperone function is its ability to promote annealing of nucleic acid strands *via* their aggregation. This activity is especially important for facilitating the bimolecular nucleation step of strand annealing. The NC-induced rate enhancement of such reactions is especially significant in solutions of low ionic strength, where it can reach up to $10^3$–$10^5$-fold (54,71–73). This is because the NC protein facilitates the new duplex nucleation by screening the negative phosphate charges of the nucleic acid strands, just as is observed in high salt. However, the facilitating effect of HIV-1 NC on duplex nucleation goes far beyond this high salt effect (72). The additional rate enhancement comes from NC-induced nucleic acid–nucleic acid attraction (74,75).

This electrostatic attraction is a phenomenon similar to the nucleic acid aggregation induced by multivalent cations (76–78). The attraction induces nonspecific aggregation of all nucleic acids, increasing the local concentration of strands, thus facilitating the diffusional search of the complementary regions for each other. The most important condition for this mechanism of nucleic acid annealing rate enhancement is that the nucleic acid molecules should remain highly mobile relative to each other within such an aggregate. These conditions seem to be fulfilled in the case of HIV-1 NCp7. Being a highly charged cationic molecule, NCp7 most likely produces the nucleic acid aggregation in a manner similar to that of other simple multivalent cations (76–82) *via* its electrostatic interaction with nucleic acids. This highly mobile and nonspecific nucleic acid aggregation induced by NCp7 can be contrasted with the much stronger and much less mobile 'sticking' of nucleic acid molecules produced by the cooperative association of the hydrophobic parts of the nucleic acid binding proteins, such as Gag. This concept is reflected in the stretching and relaxation curves shown in Figure 5, which in the cases of NCp7, NCp15, NCp9 and Gag proteins show progressively larger hysteresis.

There are several factors that may cause the hysteresis to increase in the presence of typical DNA binding proteins. First of all, a protein that binds preferentially to ssDNA, as NC does, may be slow to unbind, and therefore might interfere with strand annealing (40,83). Moreover, in order for the relaxation curve to coincide with the stretching curve, the fraction of the base pairs annealed should be in true equilibrium. This means that on the time scale of a single step, i.e. ~1 s, in which the force is being averaged in our experiment, multiple events of melting and reannealing of the base pairs should occur, such that the fraction of the annealing and melting base pairs averages to the thermodynamic equilibrium quantity. This, in turn, implies that many protein molecules bound to the ~1000 bp long piece of DNA that is melted in each step should be able to unbind ssDNA and rebind dsDNA many times on the time scale of 1 s. This is a rather stringent requirement on the ability of the protein to adjust to the state of DNA. Based on the experimental data showing very small hysteresis in the presence of NCp7, we argue here that NCp7 is indeed such a very 'fast' binding and unbinding protein.

This hypothesis is in good agreement with the very high mobility of NCp7 when bound to various nucleic acid molecules, as observed in NMR experiments. Indeed, NMR studies have shown that the majority of NCp7 molecules, when bound to nucleic acids, are able to rapidly switch between numerous possible conformations, which are averaged out on the millisecond time scale of an NMR experiment (84–86). Also, recent fluorescent energy transfer studies of the TAR DNA molecule have shown that the NCp7 (12–55) fragment is able to facilitate the rate of the terminal DNA duplex opening and closing on the $10^{-4}$–$10^{-3}$ s time scale by several-fold (87–89). This implies that NCp7, lacking the basic N-terminal tail, is capable of changing its binding mode between the ssDNA- and dsDNA-bound modes even faster than that. This is not impossible, since in order to adjust to a new state of DNA, the protein does not have to completely unbind and rebind the nucleic acid. Instead, it may simply change the conformation of its aromatic residues between DNA bases, i.e. stacked and unstacked, in addition to undergoing other small conformational changes. Such movements can easily occur on the microsecond time scale.

It is likely that protein–protein interactions resulting in cooperative protein binding to DNA are responsible for the observed hysteresis in the presence of Gag. The Gag–Gag interactions estimated from the high value of the Gag–DNA binding constant above are consistent with the large hysteresis in the presence of high concentrations of Gag (1–10 nM). Similar features are present for NCp9 and NCp15 proteins, which induce hysteresis that is much stronger than in the case of NCp7. This finding correlates with the results of Khan and Giedroc (72), who reported that, in contrast to the completely processed NCp7, NCp9 shows moderate DNA binding cooperativity. In addition, these authors found that while NCp7 binds and unbinds DNA very fast on the time scale of a Trp fluorescence quenching measurement, the NCp9 protein was shown to undergo some slow changes in its binding to DNA (72). Taken together, these findings suggest that the distal domains in the partially processed NC protein and, particularly, the unstructured, flexible and highly hydrophobic p1 sequence present in NCp9, may participate in protein–protein contacts.

Thus, the hysteresis and possible DNA–protein binding cooperativity that are observed most strongly for Gag and

NCp9 are signatures of a 'slow' DNA binding protein. Despite this slow binding characteristic, the capability of Gag to aggregate nucleic acids, most likely makes it a reasonably good chaperone for processes that do not require significant duplex destabilization, such as tRNA annealing (51–53). However, processes that require significant duplex destabilization, such as strand transfer, will likely be optimally facilitated by the final processed NCp7, which exhibits strong duplex destabilization, and which facilitates rapid reannealing of DNA that has been melted by force.

The direct results of the studies presented here can be summarized as follows. Both single DNA molecule stretching experiments and bulk Trp fluorescence assays showed that the three Gag cleavage products NCp15, NCp9 and NCp7 bound DNA with comparable affinity. The FA assays showed that NC and Gag bind with similar affinity to short DNA and RNA oligonucleotides. In contrast, the DNA stretching experiments showed that Gag bound single polymeric DNA molecules with an apparent affinity 10 times higher than that of any of the Gag cleavage products. Based on the change in single DNA molecule melting force observed in the presence of these proteins, we also found that Gag was unable to destabilize the DNA helix, while the Gag cleavage products destabilized DNA significantly. Finally, upon relaxation of the DNA after force-induced melting, we found that the amount of hysteresis was greatest for Gag, followed by NCp9 and NCp15. The final NC Gag cleavage product, NCp7, exhibited the least amount of hysteresis, indicating that this protein is most efficient at reannealing the regions of a single DNA molecule that have been melted by force. Thus, the final Gag cleavage product is expected to be most efficient nucleic acid chaperone for processes that require both nucleic acid destabilization and nucleic acid annealing. Both of these characteristics are required for minus-strand transfer, in which the very stable nucleic acid secondary structures, TAR RNA and cTAR DNA, must be annealed to each other. In contrast, Gag is more efficient at binding and packaging viral RNA. It can also perform some nucleic acid annealing functions, such as tRNA annealing, a process that does not require significant duplex destabilization, but involves nucleic acid aggregation by protein.

## ACKNOWLEDGEMENTS

This work was funded in part by the National Cancer Institute under contract No. NO1-CO-12400 with SAIC-Frederick, Inc., by the National Science Foundation (MCB-0238190), the National Institutes of Health (GM-072462), by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, and the Research Corporation. M.A.U. was partially funded by a fellowship for Perfeccionamiento y Movilidad del Personal Investigador from the Departamento de Educación, Universidades e Investigación del Gobierno Vasco (Spain). We thank Prof. Karin Musier-Forsyth for helpful discussions and Prof. Richard L. Karpel for a critical reading of the manuscript. Funding to pay the Open Access publication charges for this article was provided by NIH GM-072462.

*Conflict of interest statement*. None declared.

## REFERENCES

1. Swanstrom,R. and Wills,J. (1997) In Coffin,J. M., Hughes,S. H. and Varmus,H. (eds), *Retroviruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 263–334.
2. Campbell,S. and Rein,A. (1999) *In vitro* assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. *J. Virol.*, **73**, 2270–2279.
3. Gross,I., Hohenberg,H., Wilk,T., Wiegers,K., Grattinger,M., Muller,B., Fuller,S. and Krausslich,H.G. (2000) A conformational switch controlling HIV-1 morphogenesis. *EMBO J.*, **19**, 103–113.
4. Campbell,S., Fisher,R.J., Towler,E.M., Fox,S., Issaq,H.J., Wolfe,T., Phillips,L.R. and Rein,A. (2001) Modulation of HIV-like particle assembly *in vitro* by inositol phosphates. *Proc. Natl Acad. Sci. USA*, **98**, 10875–10879.
5. Wilk,T., Gross,I., Gowen,B.E., Rutten,T., de Haas,F., Welker,R., Krausslich,H.G., Boulanger,P. and Fuller,S.D. (2001) Organization of immature human immunodeficiency virus type 1. *J. Virol.*, **75**, 759–771.
6. Campbell,S. and Vogt,V.M. (1995) Self-assembly *in vitro* of purified CA–NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1. *J. Virol.*, **69**, 6487–6497.
7. Muriaux,D., Mirro,J., Harvin,D. and Rein,A. (2001) RNA is a structural element in retrovirus particles. *Proc. Natl Acad. Sci. USA*, **98**, 5246–5251.
8. Levin,J.G., Grimley,P.M., Ramseur,J.M. and Berezesky,I.K. (1974) Deficiency of 60 to 70S RNA in murine leukemia virus particles assembled in cells treated with actinomycin D. *J. Virol.*, **14**, 152–161.
9. Berkowitz,R., Fisher,J. and Goff,S.P. (1996) RNA packaging. *Curr. Top. Microbiol. Immunol.*, **214**, 177–218.
10. Poon,D.T., Li,G. and Aldovini,A. (1998) Nucleocapsid and matrix protein contributions to selective human immunodeficiency virus type 1 genomic RNA packaging. *J. Virol.*, **72**, 1983–1993.
11. Gorelick,R.J., Henderson,L.E., Hanser,J.P. and Rein,A. (1988) Point mutants of Moloney murine leukemia virus that fail to package viral RNA: evidence for specific RNA recognition by a 'zinc finger-like' protein sequence. *Proc. Natl Acad. Sci. USA*, **85**, 8420–8424.
12. Gorelick,R.J., Nigida,S.M.,Jr, Bess,J.W.,Jr, Arthur,L.O., Henderson,L.E. and Rein,A. (1990) Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *J. Virol.*, **64**, 3207–3211.
13. Gorelick,R.J., Chabot,D.J., Rein,A., Henderson,L.E. and Arthur,L.O. (1993) The two zinc fingers in the human immunodeficiency virus type 1 nucleocapsid protein are not functionally equivalent. *J. Virol.*, **67**, 4027–4036.
14. Jacks,T., Power,M., Masiarz,F., Luciw,P., Barr,P. and Varmus,H. (1988) Characterization of ribosomal frameshifting in HIV-1 gag–pol expression. *Nature*, **331**, 280–283.
15. Chen,N., Moragb,A., Almog,N., Blumenzweig,I., Dreazin,O. and Kotler,M. (2001) Extended nucleocapsid protein is cleaved from the Gag–Pol precursor of human immunodeficiency virus type 1. *J. Gen. Virol.*, **82**, 581–590.
16. Kräusslich,H.-G., Fäcke,M., Heuser,A.M., Konvalinka,J. and Zentgraf,H. (1995) The spacer peptide between human immunodeficiency virus capsid and nucleocapsid proteins is essential for ordered assembly and viral infectivity. *J. Virol.*, **69**, 3407–3419.
17. Pettit,S.C., Moody,M.D., Wehbie,R.S., Kaplan,A.H., Nantermet,P.V., Klein,C.A. and Swanstrom,R. (1994) The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions. *J. Virol.*, **68**, 8017–8027.
18. Shehu-Xhilaga,M., Krausslich,H.G., Pettit,S., Swanstrom,R., Lee,J.Y., Marshall,J.A., Crowe,S.M. and Mak,J. (2001) Proteolytic processing of the p2/nucleocapsid cleavage site is critical for human immunodeficiency virus type 1 RNA dimer maturation. *J. Virol.*, **75**, 9156–9164.
19. Henderson,L.E., Bowers,M.A., Sowder,R.C.,II, Serabyn,S.A., Johnson,D.G., Bess,J.W.,Jr, Arthur,L.O., Bryant,D.K. and Fenselau,C. (1992) Gag proteins of the highly replicative MN strain of human immunodeficiency virus type 1: posttranslational modifications, proteolytic processings, and complete amino acid sequences. *J. Virol.*, **66**, 1856–1865.
20. Fuller,S.D., Wilk,T., Gowen,B.E., Krausslich,H.G. and Vogt,V.M. (1997) Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. *Curr. Biol.*, **7**, 729–738.

21. Ganser,B.K., Li,S., Klishko,V.Y., Finch,J.T. and Sundquist,W.I. (1999) Assembly and analysis of conical models for the HIV-1 core. *Science*, **183**, 80–83.
22. Briggs,J.A.G., Wilk,T., Welker,R., Kräusslich,H.-G. and Fuller,S.D. (2003) Structural organization of authentic, mature HIV-1 virions and cores. *EMBO J.*, **22**, 1707–1715.
23. Briggs,J.A., Simon,M.N., Gross,I., Krausslich,H.G., Fuller,S.D., Vogt,V.M. and Johnson,M.C. (2004) The stoichiometry of Gag protein in HIV-1. *Nature Struct. Mol. Biol.*, **11**, 672–675.
24. Roldan,A., Warren,O.U., Russell,R.S., Liang,C. and Wainberg,M.A. (2005) A HIV-1 minimal Gag protein is superior to nucleocapsid at *in vitro* annealing and exhibits multimerization-induced inhibition of reverse transcription. *J. Biol. Chem.*, **280**, 17488–17496.
25. Guo,J., Wu,T., Anderson,J., Kane,B.F., Johnson,D.G., Gorelick,R.J., Henderson,L.E. and Levin,J.G. (2000) Zinc finger structures in the human immunodeficiency virus type 1 nucleocapsid protein facilitate efficient minus- and plus-strand transfer. *J. Virol.*, **74**, 8980–8988.
26. Guo,J., Wu,T., Kane,B.F., Johnson,D.G., Henderson,L.E., Gorelick,R.J. and Levin,J.G. (2002) Subtle alterations of the native zinc finger structures have dramatic effects on the nucleic acid chaperone activity of human immunodeficiency virus type 1 nucleocapsid protein. *J. Virol.*, **76**, 4370–4378.
27. Carteau,S., Gorelick,R.J. and Bushman,F.D. (1999) Coupled integration of human immunodeficiency virus type 1 cDNA ends by purified integrase *in vitro*: stimulation by the viral nucleocapsid protein. *J. Virol.*, **73**, 6670–6679.
28. Gao,K., Gorelick,R.J., Johnson,D.G. and Bushman,F.D. (2003) Cofactors for human immunodeficiency virus type 1 cDNA integration *in vitro*. *J. Virol.*, **77**, 1598–1603.
29. Kelly,R.C., Jensen,D.E. and von Hippel,P.H. (1976) DNA 'melting' proteins. IV. Fluorescence measurements of binding parameters for bacteriophage T4 gene 32-protein to mono-, oligo-, and polynucleotides. *J. Biol. Chem.*, **251**, 7240–7250.
30. McGhee,J.D. and von Hippel,P.H. (1974) Theoretical aspects of DNA–protein interactions: cooperative and non-cooperative binding of large ligands to a one-dimensional homogeneous lattice. *J. Mol. Biol.*, **86**, 469–489.
31. Record,M.T.J., Lohman,T.M. and de Haseth,P.L. (1976) Ion effects on ligand–nucleic acid interactions. *J. Mol. Biol.*, **107**, 145–158.
32. Record,M.T.J., Zhang,W. and Anderson,C.F. (1998) Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: a practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts. *Adv. Protein Chem.*, **51**, 281–353.
33. Perrin,F. (1926) Polarization de la luminiere de fluorescence. Vie moyenne de molecules dans l'etat excite. *J. Phys. Radium*, **7**, 390–401.
34. Lakowicz,J.R. (1999) *Principles of Fluorescence Spectroscopy*. 2nd edn. Plenum Publishers, NY.
35. Williams,M.C., Rouzina,I., Wenner,J.R., Gorelick,R.J., Musier-Forsyth,K. and Bloomfield,V.A. (2001) Mechanism for nucleic acid chaperone activity of HIV-1 nucleocapsid protein revealed by single molecule stretching. *Proc. Natl Acad. Sci. USA*, **98**, 6121–6126.
36. Williams,M.C., Gorelick,R.J. and Musier-Forsyth,K. (2002) Specific zinc finger architecture required for HIV-1 nucleocapsid protein's nucleic acid chaperone function. *Proc. Natl Acad. Sci. USA*, **99**, 8614–8619.
37. Lee,B.M., De Guzman,R.N., Turner,B.G., Tjandra,N. and Summers,M.F. (1998) Dynamical behavior of the HIV-1 nucleocapsid protein. *J. Mol. Biol.*, **279**, 633–649.
38. Martin,S.L., Cruceanu,M., Branciforte,D., Li,P.W.-l., Kwok,S.C., Hodges,R.S. and Williams,M.C. (2005) LINE-1 retrotransposition requires the nucleic acid chaperone activity of the ORF1 protein. *J. Mol. Biol.*, **348**, 549–561.
39. Priel,E., Aflalo,E., Seri,I., Henderson,L.E., Arthur,L.O., Aboud,M., Segal,S. and Blair,D.G. (1995) DNA binding properties of the zinc-bound and zinc-free HIV nucleocapsid protein: supercoiled DNA unwinding and DNA–protein cleavable complex formation. *FEBS Lett.*, **362**, 59–64.
40. Pant,K., Karpel,R.L., Rouzina,I. and Williams,M.C. (2004) Mechanical measurement of single molecule binding rates: kinetics of DNA helix-destabilization by T4 gene 32 protein. *J. Mol. Biol.*, **336**, 851–870.
41. Cantor,C.R. and Schimmel,P.R. (1980) *Biophysical Chemistry*. W. H. Freeman and Company, New York.
42. Rouzina,I. and Bloomfield,V.A. (2001) Force-induced melting of the DNA double helix-1. Thermodynamic analysis. *Biophys. J.*, **80**, 882–893.
43. Williams,M.C., Rouzina,I. and Bloomfield,V.A. (2002) Thermodynamics of DNA interactions from single molecule stretching experiments. *Acc. Chem. Res.*, **35**, 159–166.
44. Leger,J.F., Romano,G., Sarkar,A., Robert,J., Bourdieu,L., Chatenay,D. and Marko,J.F. (1999) Structural transitions of a twisted and stretched DNA molecule. *Phys. Rev. Lett.*, **83**, 1066–1069.
45. Urbaneja,M.A., Kane,B.P., Johnson,D.G., Gorelick,R.J., Henderson,L.E. and Casas-Finet,J.R. (1999) Binding properties of the human immunodeficiency virus type 1 nucleocapsid protein p7 to a model RNA: elucidation of the structural determinants for function. *J. Mol. Biol.*, **287**, 59–75.
46. Bombarda,E., Ababou,A., Vuilleumier,C., Gérard,D., Roques,B.P., Piémont,E. and Mély,Y. (1999) Time-resolved fluorescence investigation of the human immunodeficiency virus type 1 nucleocapsid protein: influence of the binding of nucleic acids. *Biophys. J.*, **76**, 1561–1570.
47. Urbaneja,M.A., Wu,M., Casas-Finet,J.R. and Karpel,R.L. (2002) HIV-1 nucleocapsid protein as a nucleic acid chaperone: spectroscopic study of its helix-destabilizing properties, structural binding specificity, and annealing activity. *J. Mol. Biol.*, **318**, 749–764.
48. Fisher,R.J., Rein,A., Fivash,M., Urbaneja,M.A., Casas-Finet,J.R., Medaglia,M. and Henderson,L.E. (1998) Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides. *J. Virol.*, **72**, 1902–1909.
49. Vuilleumier,C., Bombarda,E., Morellet,N., Gerard,D., Roques,B.P. and Mely,Y. (1999) Nucleic acid sequence discrimination by the HIV-1 nucleocapsid protein NCp7: a fluorescence study. *Biochemistry*, **38**, 16816–16825.
50. Paoletti,A.C., Shubsda,M.F., Hudson,B.S. and Borer,P.N. (2002) Affinities of the nucleocapsid protein for variants of SL3 RNA in HIV-1. *Biochemistry*, **41**, 15423–15428.
51. Cen,S., Huang,Y., Khorchid,A., Darlix,J.-L., Wainberg,M.A. and Kleiman,L. (1999) The role of Pr55(gag) in the annealing of tRNA3Lys to human immunodeficiency virus type 1 genomic RNA. *J. Virol.*, **73**, 4485–4488.
52. Crawford,S. and Goff,S.P. (1985) A deletion mutation in the 5′ part of the pol gene of Moloney murine leukemia virus blocks proteolytic processing of the gag and pol polyproteins. *J. Virol.*, **53**, 899–907.
53. Feng,Y.-X., Campbell,S., Harvin,D., Ehresmann,B., Ehresmann,C. and Rein,A. (1999) The human immunodeficiency virus type 1 Gag polyprotein has nucleic acid chaperone activity: possible role in dimerization of genomic RNA and placement of tRNA on the primer binding site. *J. Virol.*, **73**, 4251–4256.
54. Hargittai,M.R.S., Gorelick,R.J., Rouzina,I. and Musier-Forsyth,K. (2004) Mechanistic insights into the kinetics of HIV-1 nucleocapsid protein-facilitated tRNA annealing to the primer binding site. *J. Mol. Biol.*, **337**, 951–968.
55. Guo,J.H., Henderson,L.E., Bess,J., Kane,B. and Levin,J.G. (1997) Human immunodeficiency virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-stop DNA. *J. Virol.*, **71**, 5178–5188.
56. Golinelli,M.P. and Hughes,S.H. (2001) Self-priming of retroviral minus-strand strong-stop DNAs. *Virology*, **285**, 278–290.
57. Heilman-Miller,S.L., Wu,T. and Levin,J.G. (2004) Alteration of nucleic acid structure and stability modulates the efficiency of minus-strand transfer mediated by the HIV-1 nucleocapsid protein. *J. Biol. Chem.*, **279**, 44154–44165.
58. Driscoll,M.D. and Hughes,S.H. (2000) Human immunodeficiency virus type 1 nucleocapsid protein can prevent self-priming of minus-strand strong stop DNA by promoting the annealing of short oligonucleotides to hairpin seqeunces. *J. Virol.*, **74**, 8785–8792.
59. Negroni,M. and Buc,H. (1999) Recombination during reverse transcription: an evaluation of the role of the nucleocapsid protein. *J. Mol. Biol.*, **286**, 15–31.
60. Negroni,M. and Buc,H. (2001) Mechanisms of retroviral recombination. *Ann. Rev. Genet.*, **35**, 275–302.
61. Chen,Y., Balakrishnan,M., Roques,B. and Bambara,R.A. (2003) Steps of the acceptor invasion mechanism for HIV-1 minus strand strong stop transfer. *J. Biol. Chem.*, **278**, 38368–38375.
62. Chen,Y., Balakrishnan,M., Roques,B., Fay,P.J. and Bambara,R.A. (2003) Mechanism of minus strand strong stop transfer in HIV-1 reverse transcription. *J. Biol. Chem.*, **278**, 8006–8017.

63. Levin,J.G., Guo,J., Rouzina,I. and Musier-Forsyth,K. (2005) Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism. *Prog. Nucleic Acid Res. Mol. Biol.*, **80**, 217–286.
64. De Guzman,R.N., Wu,Z.R., Stalling,C.C., Pappalardo,L., Borer,P.N. and Summers,M.F. (1998) Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science*, **279**, 384–388.
65. Amarasinghe,G.K., Guzman,R.N.D., Turner,R.B., Chancellor,K.J., Wu,Z.R. and Summers,M.F. (2000) NMR Structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. *J. Mol. Biol.*, **301**, 491–511.
66. Mely,Y., Piemont,E., Sorinas-Jimeno,M., De Rocquigny,H., Jullian,N., Morellet,N., Roques,B.P. and Gerard,D. (1993) Structural and dynamic characterization of the aromatic amino acids of the human immunodeficiency virus type I nucleocapsid protein zinc fingers and their involvement in heterologous tRNA(Phe) binding: a steady-state and time-resolved fluorescence study. *Biophys. J.*, **65**, 1513–1522.
67. Li,S., Hill,C.P., Sundquist,W.I. and Finch,J.T. (2000) Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature*, **407**, 409–413.
68. Mayo,K., Huseby,D., McDermott,J., Arvidson,B., Finlay,L. and Barklis,E. (2003) Retrovirus capsid protein assembly arrangements. *J. Mol. Biol.*, **325**, 225–237.
69. Ehrlich,L.S., Agresta,B.E. and Carter,C.A. (1992) Assembly of recombinant human immunodeficiency virus type 1 capsid protein *in vitro*. *J. Virol.*, **66**, 4874–4883.
70. Zlotnick,A. (2003) Are weak protein–protein interactions the general rule in capsid assembly? *Virology*, **315**, 269–274.
71. You,J.C. and McHenry,C.S. (1994) Human immunodeficiency virus nucleocapsid protein accelerates strand transfer of the terminally redundant sequences involved in reverse transcription. *J. Biol. Chem.*, **269**, 31491–31495.
72. Khan,R. and Giedroc,D.P. (1994) Nucleic acid binding properties of recombinant Zn2 HIV-1 nucleocapsid protein are modulated by COOH-terminal processing. *J. Biol. Chem.*, **269**, 22538–22546.
73. Tsuchihashi,Z. and Brown,P.O. (1994) DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein. *J. Virol.*, **68**, 5863–5870.
74. Le Cam,E., Coulaud,D., Delain,E., Petitjean,P., Roques,B.P., Gérard,D., Stoylova,E., Vuilleumier,C., Stoylov,S.P. and Mély,Y. (1998) Properties and growth mechanism of the ordered aggregation of a model RNA by the HIV-1 nucleocapsid protein: an electron microscopy investigation. *Biopolymers*, **45**, 217–229.
75. Stoylov,S.P., Vuilleumier,C., Stoylova,E., Rocquigny,H.D., Roques,B.P., Gérard,D. and Mély,Y. (1997) Ordered aggregation of ribonucleic acids by the human immunodeficiency virus type 1 nucleocapsid protein. *Biopolymers*, **41**, 301–312.
76. Bloomfield,V.A. (1997) DNA condensation by multivalent cations. *Biopolymers*, **44**, 269–282.
77. Krishnamoorthy,G., Roques,B., Darlix,J.-L. and Mely,Y. (2003) DNA condensation by the nucleocapsid protein of HIV-1: a mechanism ensuring DNA protection. *Nucleic Acids Res.*, **31**, 5425–5432.
78. Lai,E. and van Zanten,J.H. (2001) Monitoring DNA/poly-L-lysine polyplex formation with time-resolved multiangle laser light scattering. *Biophys. J.*, **80**, 864–873.
79. Nguyen,T.T., Rouzina,I. and Shklovskii,B.I. (2000) Reentrant condensation of DNA induced by multivalent counterions. *J. Chem. Phys.*, **112**, 2562–2568.
80. Rouzina,I. and Bloomfield,V.A. (1996) Macroion attraction due to electrostatic correlation between screening counterions. 1. Mobile surface-adsorbed ions and diffuse ion cloud. *J. Phys. Chem.*, **100**, 9977–9989.
81. Koculi,E., Lee,N.K., Thirumalai,D. and Woodson,S.A. (2004) Folding of the Tetrahymena ribozyme by polyamines: importance of counterion valence and size. *J. Mol. Biol.*, **341**, 27–36.
82. Heilman-Miller,S.L., Pan,J., Thirumalai,D. and Woodson,S.A. (2001) Role of counterion condensation in folding of the Tetrahymena ribozyme. II. Counterion-dependence of folding kinetics. *J. Mol. Biol.*, **309**, 57–68.
83. Pant,K., Karpel,R.L. and Williams,M.C. (2003) Kinetic regulation of single DNA molecule denaturation by T4 gene 32 protein structural domains. *J. Mol. Biol.*, **327**, 571–578.
84. Johnson,P.E., Turner,R.B., Wu,Z.R., Hairston,L., Guo,J., Levin,J.G. and Summers,M.F. (2000) A mechanism for plus-strand transfer enhancement by the HIV-1 nucleocapsid protein during reverse transcription. *Biochemistry*, **39**, 9084–9091.
85. Tisne,C., Roques,B. and Dardel,F. (2001) Heteronuclear NMR studies of the interaction of tRNA(Lys)3 with HIV-1 nucleocapsid protein. *J. Mol. Biol.*, **306**, 443–454.
86. Tisne,C., Roques,B.P. and Dardel,F. (2003) Specific recognition of primer tRNA Lys 3 by HIV-1 nucleocapsid protein: involvement of the zinc fingers and the N-terminal basic extension. *Biochimie*, **85**, 557–561.
87. Azoulay,J., Clamme,J.-P., Darlix,J.L., Roques,B.P. and Mely,Y. (2003) Destabilization of the HIV-1 complementary sequence of TAR by the nucleocapsid protein through activation of conformational fluctuations. *J. Mol. Biol.*, **326**, 691–700.
88. Bernacchi,S., Stoylov,S.P., Piemont,E., Ficheux,D., Roques,B., Darlix,J.L. and Mely,Y. (2002) HIV-1 nucleocapsid protein activates transient melting of least stable parts of the secondary structure of TAR and its complementary sequence. *J. Mol. Biol.*, **317**, 385–399.
89. Beltz,H., Azoulay,J., Bernacchi,S., Clamme,J.P., Ficheux,D., Roques,B., Darlix,J.-L. and Mely,Y. (2003) Impact of the terminal bulges of HIV-1 cTAR DNA on its stability and the destabilizing activity of the nucleocapsid protein NCp7. *J. Mol. Biol.*, **328**, 95–108.
90. Smith,S.B., Cui,Y.J. and Bustamante,C. (1996) Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science*, **271**, 795–799.