Published online 14 September 2011 *Nucleic Acids Research, 2012, Vol. 40, No. 2* ***751–760***
doi:10.1093/nar/gkr726

A single zinc finger optimizes the DNA interactions of the nucleocapsid protein of the yeast retrotransposon Ty3

Kathy R. Chaurasiya[1], Hylkje Geertsema[1], Gaël Cristofari[2], Jean-Luc Darlix[2] and Mark C. Williams[1,3,*]

[1]Department of Physics, Northeastern University, Boston, MA, USA, [2]Unité de Virologie Humaine INSERM 758, IFR 128 Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69364 Lyon, France and [3]Center for Interdisciplinary Research on Complex Systems, Northeastern University, Boston, MA, USA

Received June 22, 2011; Revised August 9, 2011; Accepted August 23, 2011

ABSTRACT

Reverse transcription in retroviruses and retrotransposons requires nucleic acid chaperones, which drive the rearrangement of nucleic acid conformation. The nucleic acid chaperone properties of the human immunodeficiency virus type-1 (HIV-1) nucleocapsid (NC) protein have been extensively studied, and nucleic acid aggregation, duplex destabilization and rapid binding kinetics have been identified as major components of its activity. However, the properties of other nucleic acid chaperone proteins, such as retrotransposon Ty3 NC, a likely ancestor of HIV-1 NC, are not well understood. In addition, it is unclear whether a single zinc finger is sufficient to optimize the properties characteristic of HIV-1 NC. We used single-molecule DNA stretching as a method for detailed characterization of Ty3 NC chaperone activity. We found that wild type Ty3 NC aggregates single- and double-stranded DNA, weakly stabilizes dsDNA, and exhibits rapid binding kinetics. Single-molecule studies in the presence of Ty3 NC mutants show that the N-terminal basic residues and the unique zinc finger at the C-terminus are required for optimum chaperone activity in this system. While the single zinc finger is capable of optimizing Ty3 NC's DNA interaction kinetics, two zinc fingers may be necessary in order to facilitate the DNA destabilization exhibited by HIV-1 NC.

INTRODUCTION

Retrotransposons are mobile genetic elements, closely related to retroviruses, which use their own proteins in conjunction with cellular machinery to replicate independently of the genome. The yeast retrotransposon Ty3, and retroviruses such as human immunodeficiency virus type-1 (HIV-1), encode reverse transcriptase (RT), a DNA polymerase that converts the single-stranded RNA (ssRNA) genome of positive polarity into double-stranded DNA (dsDNA). During this reverse transcription process, two obligatory DNA strand transfers occur to generate the long terminal repeats (LTR) that control genomic DNA integration and its expression (1). In the early stages of replication, a tRNA primer anneals to the primer binding site (PBS) close to the 5′-end of the ssRNA, where RT synthesizes the complementary DNA (cDNA) strand, referred to as minus strand strong stop DNA (-sssDNA). While RT synthesizes -ssDNA, its RNaseH domain degrades the 5′-end of the RNA template. Further DNA polymerization requires -sssDNA transfer, where the newly synthesized single-stranded cDNA anneals to the 3′-untranslated terminal region (UTR) of the RNA template (1–3). However, thermodynamically stable structures within the cDNA and the 3′-UTR RNA make duplex formation improbable. The nucleocapsid (NC) protein acts as a nucleic acid chaperone that destabilizes these secondary structures and facilitates DNA–RNA hybrid formation (4–8). NC also participates in the second DNA strand transfer, where the plus DNA strand anneals to the structured PBS region of the minus DNA strand (9–12). These transfer events are required for RT to transcribe the remaining genomic RNA into dsDNA flanked by the LTR, which the homologous enzyme integrase inserts into the cellular genome.

Replication of retroelements requires the nucleic acid chaperone activity of NC proteins, which tend to be small cationic proteins with minimal structure other than one or two CCHC-motif zinc fingers (1–3,13–17). HIV-1 NC is 55 amino acids in length, with two zinc fingers and a

*To whom correspondence should be addressed. Tel: +1 617 373 7323; Fax: +1 617 373 2943; Email: mark@neu.edu

basic N-terminal tail (1–3). It directs annealing of the $tRNA_i^{Lys}$ primer to the PBS of the genomic RNA (5,18,19). The viral RNA has 5′ and 3′ repeat regions that include stable *trans*-activation response element (TAR) hairpins, and HIV-1 NC facilitates rearrangement of these nucleic acids during the minus strand transfer step of reverse transcription (10–12). NC is also an essential nucleic acid chaperone involved in genomic RNA dimerization and virus assembly (20–24). The chaperone activity of HIV-1 NC involves nucleic acid aggregation, duplex destabilization and rapid binding kinetics (25–30). Nucleic acid aggregation, an effect due to protein-induced interactions that make the DNA molecule attracted to itself, is associated with the cationic domain of HIV-1 NC. Duplex destabilization involves aromatic residues in the zinc fingers, which stack with ssDNA bases (31,32). Thus optimal chaperone activity of HIV-1 NC is highly sensitive to the zinc-finger architecture, and relies on a delicate balance between destabilizing secondary structures and promoting complementary strand annealing.

A recent study examining the nucleic acid chaperone activity of NC proteins from several retroviruses established that the nucleic acid interaction characteristics that are important for nucleic acid chaperone activity varied significantly for different viruses (33,34). From these studies, it is clear that HIV-1 NC has optimal chaperone activity compared to other NC proteins. Although nucleic acid chaperones from HIV-1 and other retroviruses have been extensively studied, less is known about chaperone activity in LTR retrotransposons (4,35,36). Since Ty3 NC is a likely ancestor of HIV-1 NC, here we examine its nucleic acid chaperone properties and compare them to those of HIV-1 NC. In addition to allowing comparison of Ty3 NC to the paradigm of nucleic acid chaperone proteins, understanding these differences may also be useful for the development of drugs that target the NC protein from the rapidly evolving HIV-1 retrovirus (37–39). Ty3 NC is a 57-residue protein with one zinc finger and a basic N-terminal tail (40). Ty3 NC is monomeric in solution, and it is required for efficient annealing of the $tRNA_i^{Met}$ primer on the bipartite Ty3 PBS and RNA dimerization (41). The UTR regions of the Ty3 genome also form secondary structures, and minus strand transfer requires Ty3 NC to overcome the energetic barrier in forming the DNA–RNA duplex necessary for further reverse transcription.

There are distinct structural differences between Ty3 NC, which has only one zinc finger with two aromatic residues, and HIV-1 NC, which has two zinc fingers with one aromatic residue each. The cationic tail of Ty3 NC, which has minimal structure, also contains an aromatic residue (tyrosine). Ty3 NC, with 20 basic residues (pI $\approx$ 11.5), has a slightly higher overall charge than HIV-1 NC (pI $\approx$ 10), with 15 basic residues. Although minimal information is available about the 3D structure of Ty3 NC, it is extremely likely to have high charge density, consistent with retroviral NC proteins of known structure such as HIV-1 NC and Moloney murine leukemia virus (MuLV) NC (42). Furthermore, the length and structural complexity of the UTR regions seem to be related to the chaperone activity of a number of retroviral NC proteins, and the secondary structures in the Ty3 UTR regions are not as thermodynamically stable as the HIV-1 TAR hairpins (25).

DNA stretching and other biophysical methods have been used to show that the properties of HIV-1 NC are specifically tuned to optimize its interactions with both dsDNA and ssDNA. Specifically, both zinc fingers are required to maintain their native order and structure, and seemingly minor mutations result in loss of rapid kinetics of the DNA–protein interaction (26). Other retroviral NC proteins exhibited less rapid kinetics in DNA stretching experiments, particularly MuLV and human T-cell lymphotropic virus type 1 (HTLV-1) NC (33,34). Notably, MuLV NC is the only NC protein studied that contained a single zinc finger, and it exhibited relatively slow kinetics. Thus, previous results suggest that two zinc fingers are generally required to optimize the kinetics of NC–DNA interactions (26,27,33,34). To determine the capability of a single zinc finger to facilitate NC–DNA interactions and to understand the connection between Ty3 NC's structure and function as a nucleic acid chaperone, we probed the thermodynamics and kinetics of wild type and mutant NC–DNA interactions with single DNA molecules. We found that, despite the presence of only a single zinc finger, wild-type Ty3 NC aggregates nucleic acids, promotes complementary strand annealing, and exhibits rapid kinetics resembling that of HIV-1 NC. Four mutants of Ty3 NC demonstrate that these chaperone properties require the zinc finger as well as the presence of its cationic N-terminal tail. Thus, Ty3 NC's single zinc finger is necessary and sufficient to facilitate rapid DNA–protein interactions, but a second zinc finger may be required for more effective nucleic acid destabilization by NC proteins, such as that observed for HIV-1 NC and TAR hairpins.

MATERIALS AND METHODS

Single-molecule DNA stretching experiments were performed as previously described (35,43). Dual beam optical tweezers were used to capture biotin-labeled bacteriophage λ DNA between two streptavidin-coated polystyrene beads (Bangs Labs). Surrounding DNA molecules were rinsed out of solution with buffer (50 mM Na^+, 10 mM HEPES, pH 7.5), and the captured DNA molecule was then stretched and released at a pulling rate of 100 nm/s to obtain DNA-only force-extension curves (Figure 1a, black), which are typically characterized by models of polymer elasticity. The worm-like chain (WLC) model describes dsDNA in terms of length b_{ds} at a given force F:

$$b_{\mathrm{ds}}(F) = B_{ds}\left[1 - \frac{1}{2}\left(\frac{k_B T}{P_{\mathrm{ds}} F}\right)^{\frac{1}{2}} + \frac{F}{S_{\mathrm{ds}}}\right] \quad (1)$$

where P_{ds} is the persistence length, B_{ds} is the contour length and S_{ds} is the stretch modulus. The freely jointed

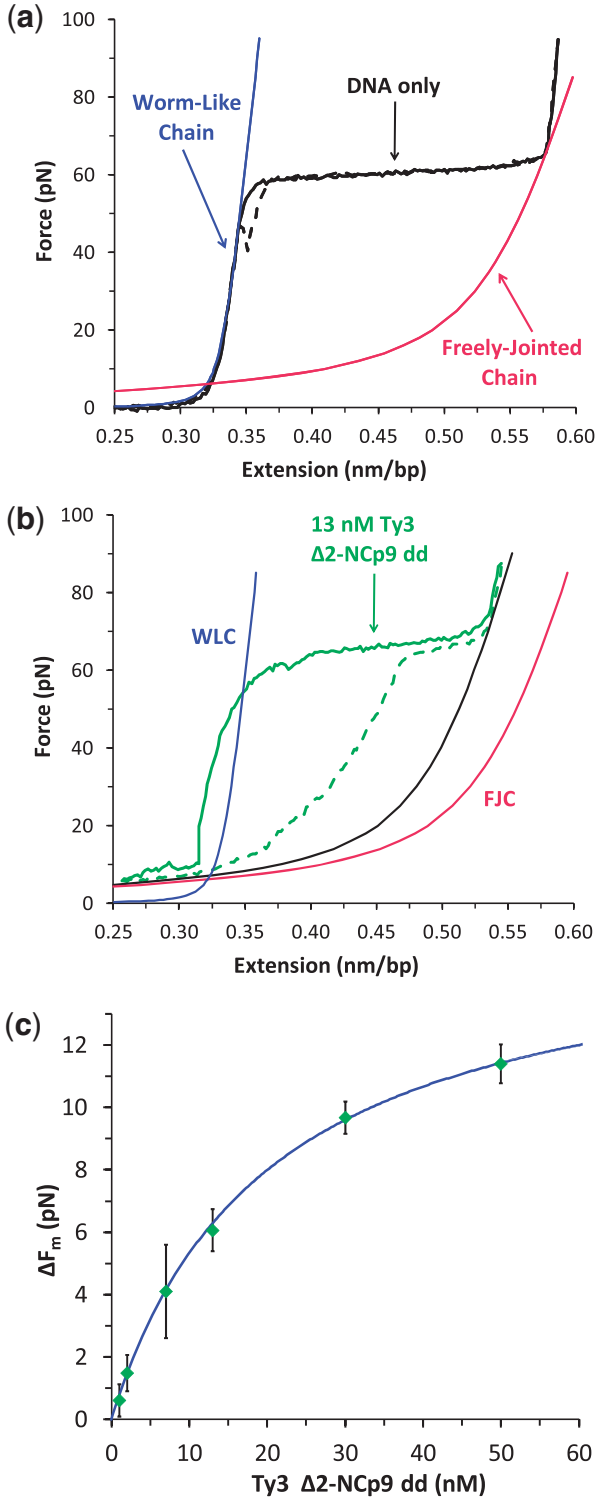

Figure 1. (**a**) Typical force-extension (solid) and release (dashed) curves of λ-DNA (black) obtained with optical tweezers. The WLC model (blue line) describes dsDNA. Near the dsDNA contour length, the molecule undergoes a force-induced melting transition, from dsDNA to ssDNA. The FJC model describes ssDNA (red line). Minimal hysteresis is evident in these solution conditions (50 mM Na^+, 10 mM HEPES, pH 7.5). (**b**) Quantification of the hysteresis area ratio for a typical DNA extension and release curve. Force–extension (solid) and release (dashed) curve of DNA in the presence of Ty3 NC Δ2-NCp9 dd are shown in green. The WLC and FJC models are shown in blue and red, respectively. A linear combination of these two models is shown in black, indicating the fraction of ssDNA exposed to solution upon DNA extension. Relative hysteresis is the ratio of A_1, the area between the stretch (solid green) and release (dashed green) curves, and A_2, the area between the stretch (solid green) and melted DNA fraction (black) curves. (**c**) Equilibrium dissociation constant K_d determined from change in average melting force ΔF_m as a function of protein concentration c fit to a simple DNA binding isotherm [Equations (3) and (5)]. Data points for mutant Δ2-NCp9 dd are shown with standard error bars, with a fit (blue line) that yields $K_d = 20\ (\pm\ 1)$ nM and saturated melting force $\Delta F_m^{sat} = 16\ (\pm\ 0.5)$ pN. K_d was estimated for mutants Δ1-NCp9 and NCp9 dd with this method, but could not be obtained for mutant Δ2-NCp9, which did not affect DNA melting force.

chain (FJC) model describes ssDNA in terms of length b_{ss} at a given force F:

$$b_{ss}(F) = B_{ss}\left[\coth\left(\frac{2P_{ss}F}{k_BT}\right) - \frac{1}{2}\frac{k_BT}{P_{ss}F}\right]\left[1+\frac{F}{S_{ss}}\right]. \quad (2)$$

Figure 1a shows the WLC (blue) and FJC (red) polymer models with typical parameter values [$B_{ds} = 0.34$ nm/bp, $P_{ds} = 48$ nm and $S_{ds} = 1200$ pN in Equation (1), $B_{ss} = 0.55$ nm/bp, $P_{ss} = 0.75$ nm and $S_{ss} = 720$ pN in Equation (2)]. After obtaining force-extension curves of a single DNA molecule, the buffer was replaced with protein dilutions. Ty3 NC and its mutants were introduced into solution with the DNA held under slight tension in an effort to reduce formation of extremely stable protein–DNA aggregates during solution exchange. The first force-extension curve after protein exchange was therefore stretched from zero extension to the original starting extension after DNA release (example in Figure 1b, solid green curve, starting force of 10 pN). Subsequent force-extension curves begin and end at zero extension. At least three stretch–release cycles were performed for each measurement, and subsequent force-extension curves were the same within error in the case of wild type Ty3 NC. The 50 mM Na^+ concentration was used for comparison with earlier studies of other NC proteins. Both the salt and protein concentration *in vivo* are higher than those used here, but the increased binding in low salt tends to compensate for the lower protein concentration.

Melting forces were determined by averaging along the length of the force-induced melting plateau in the presence of each protein. Hysteresis was quantified by calculating A_1, the area between the force-extension (Figure 1b, solid green) and release (Figure 1b, dashed green) curves of DNA in the presence of wild-type Ty3 NC and its mutants. A linear combination of the WLC and FJC models (Figure 1b, blue and red, respectively) indicates the fraction of ssDNA generated upon DNA extension (Figure 1b, black), as described in (44). The total amount of hysteresis possible is A_2, the area between the extension curve (Figure 1b, solid green) and this linear combination (Figure 1b, black). Relative hysteresis is the ratio of A_1 and A_2, ranging from 0.1 (±0.03) in the absence of protein, to a theoretical maximum of 1, which would indicate that all the ssDNA generated is bound by protein upon DNA release and does not dissociate on the time scale of the release (~1 min). Averages and uncertainties (standard

error) for all values reported were calculated using at least three measurements.

To quantify the effect of wild-type Ty3 NC on the shape of the force-induced melting plateau, transition width δF was determined from the slope at the midpoint of melting transition, as described in (27,45). The change in transition width is $\Delta F = \delta F - \delta F_0$, where δF_0 is the transition width in the absence of protein. Change in transition width as a function of protein concentration c may be described by a simple DNA binding isotherm (46):

$$\theta = \frac{c}{K_d}(1-\theta) \quad (3)$$

where K_d is the equilibrium dissociation constant and the fractional DNA binding θ is given by the change in transition width:

$$\theta = \frac{\Delta F(c)}{\Delta F_{sat}} = \frac{\delta F(c) - \delta F_0}{\delta F_{sat} - \delta F_0} \quad (4)$$

with δF_{sat} defined as the transition width at saturated protein concentration. Standard error for measured data points was calculated from at least three measurements, and fits to Equations (3) and (4) were performed using χ^2 analysis.

Ty3 NC mutants have negligible effect on the transition width, so K_d was obtained from their effect on the average melting force F_m, as reviewed in (43). Change in melting force is $\Delta F_m = F_m - F_m^0$, where $F_m^0 = 61.0$ (0.5) pN, the melting force in the absence of protein. ΔF_m as a function of protein concentration c may be described by Equation (3), where the fractional binding θ is given by the change in melting force:

$$\theta = \frac{\Delta F_m(c)}{\Delta F_m^{sat}} = \frac{F_m(c) - F_m^0}{F_m^{sat} - F_m^0} \quad (5)$$

with F_m^{sat} defined as the average melting force at saturated protein concentration. Data points with standard error were fit with Equations (3) and (5) to quantify K_d for mutant Δ2-NCp9 dd (Figure 1c). This method was used to estimate K_d for mutants Δ1-NCp9 and NCp9 dd, but mutant Δ2-NCp9 did not affect melting force over two orders of magnitude change in protein concentration.

Ty3 NC protein of 57 amino acids and its mutants were synthesized on the solid phase using fmoc (9-fluorenylmethoxycarbonyl)- and opfp (pentafluorophenyl ester)-protected amino acids and purified to homogeneity by HPLC (47).

RESULTS

Wild-type Ty3 NC binds DNA with rapid kinetics and slightly stabilizes dsDNA

Optical tweezers were used to stretch and release a single molecule of bacteriophage λ DNA in the absence of protein and in the presence of a nearly saturated concentration of wild type Ty3 NC (Figure 2a). While bacteriophage λ DNA is not the sequence acted on by Ty3 NC, it

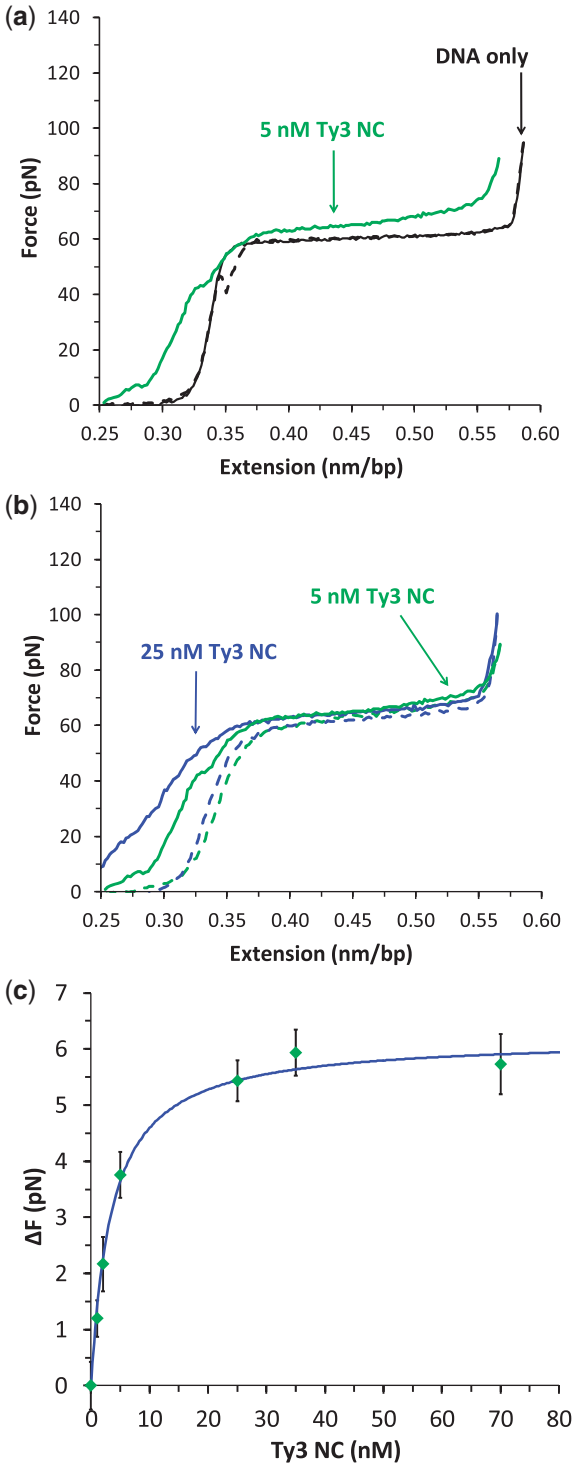

Figure 2. (**a** and **b**) Typical force-extension (solid) and release (dashed) curves of λ-DNA in the presence of wild type Ty3 NC. (**a**) DNA only (black) and 5 nM wild type Ty3 NC (green). (**b**) 5 nM (green) and 25 nM (blue) wild type Ty3 NC. (**c**) Change in the transition width ΔF of DNA force-induced melting as a function of wild type Ty3 NC concentration. $\Delta F = \delta F - \delta F_0$, where δF is the melting transition width in the presence of protein and $\delta F_0 = 3.6$ (±0.3) pN, the melting transition width of DNA only. Standard error determined from at least three measurements was used to compute error bars for ΔF. A χ^2 fit (blue line) to a simple DNA binding isotherm [Equations (3) and (4)] yields $K_d = 3.5$ (±0.5) nM and $\Delta F_{sat} = 6.2$ (±0.4) pN. Protein concentrations significantly above saturation (80–150 nM) were also included in the χ^2 fit (data not shown).

represents a random sequence, allowing us to measure the ability of the protein to remodel general nucleic acid structures. In the absence of protein, B-form dsDNA uncoils with a gradual increase in force at low extensions. As the molecule approaches its contour length of 0.34 nm/bp, the elasticity of the backbone causes a sharp increase in force. At ~60 pN, the DNA molecule lengthens more than 1.5 times with minimal increase in force as it undergoes a cooperative force-induced melting transition from dsDNA to ssDNA (48–50). The force increases rapidly again at the end of this cooperative phase transition, at 0.6 nm/bp. As the DNA is released back to low extensions, the force-extension curve is almost completely reversible, exhibiting only minor hysteresis, which is the difference between the extension and release curves. The area between these curves characterizes the amount of hysteresis.

In the presence of 5 nM wild type Ty3 NC, the force increases at low extensions, below the λ DNA contour length (Figure 2a). The transition at the end of the melting plateau also shifts to lower extensions, decreasing by ~0.025 nm/bp. These effects are due to nucleic acid aggregation, in which protein-induced interactions make the DNA molecule attracted to itself. A 5-fold increase in protein concentration leads to small increases in these two effects, reflecting additional nucleic acid aggregation (Figure 2b).

The force-induced melting plateau is slightly sloped in the presence of wild type Ty3 NC. This increase in transition width reflects a loss of DNA melting cooperativity, which indicates that the DNA molecule is more likely to undergo conformational rearrangements, an effect which has been observed for multiple nucleic acid chaperone proteins (27,34,45,46). A simple DNA binding isotherm [Equations (3) and (4)] fit to the change in transition width ΔF as a function of protein concentration c yields an equilibrium dissociation constant $K_d = 3.5\ (\pm 0.5)$ nM and saturated change in transition width $\Delta F_{sat} = 6.2\ (\pm 0.4)$ pN (Figure 2c, blue line).

The lack of hysteresis indicates that Ty3 NC exhibits rapid kinetics, allowing it to dissociate quickly from ssDNA. Rapid kinetics, which is characteristic of nucleic acid chaperones, is correlated with a protein's ability to facilitate re-annealing. In contrast, proteins that preferentially bind ssDNA, such as T4 gp32, induce significant hysteresis, and proteins that preferentially bind dsDNA, such as HMG, stabilize the duplex and increase the melting force (43). The melting force in the presence of wild type Ty3 NC is slightly higher than that of DNA only (Table 1), which indicates net stabilization of the DNA duplex. In order to elucidate how a protein that stabilizes duplex DNA can act as a nucleic acid chaperone, four mutants of Ty3 NC were used to further investigate the DNA binding role of the zinc finger and N-terminal tail (Figure 3). The N-terminal tail was deleted to varying degrees in mutants Δ1-NCp9 and Δ2-NCp9, leaving the zinc finger intact. In mutant Δ2-NCp9 dd, however, the zinc finger has also been deleted. Mutant NCp9 dd has a largely intact N-terminal tail, but no zinc finger.

Table 1. Melting force and hysteresis area in the presence of wild type Ty3 NC and its mutants, measured at protein concentrations c (force-extension curves shown in Figures 2, 4–7), near or above K_d. Equilibrium dissociation constants K_d reported for 50 nM Na^+

Ty3 NC protein	c (nM)	F_m (pN)[a]	Hysteresis (ratio)[b]	K_d (nM)[c]
Wild type	5	68.0 (±1.0)	0.26 (±0.03)	3.5 (±0.5)
Δ1-NCp9	20	59.4 (±0.7)	0.68 (±0.1)	12 (±3)
Δ2-NCp9[d]	50	60.6 (±0.1)	0.49 (±0.1)	
NCp9 dd	3	71.5 (±0.4)	0.55 (±0.1)	3 (±2)
Δ2-NCp9 dd	13	67.1 (±0.5)	0.67 (±0.1)	20 (±1)

All values were calculated with at least three measurements, reflecting the uncertainty reported (standard error).
[a]The melting force is an average along the length of the force-induced melting plateau in the presence of each protein. The melting force of DNA only in 50 mM Na^+ is 61.0 (±0.5) pN.
[b]Hysteresis is an area ratio that reflects amount of protein still bound upon DNA release (Figure 1b). The ratio increases with the amount of ssDNA observed upon DNA release, up to a maximum value of 1. The minimal hysteresis in the absence of protein yields an area ratio of 0.1 (±0.03).
[c]K_d was quantified from the change in transition width in the case of wild-type Ty3 NC (Figure 2c). Change in melting force was used to determine K_d for Ty3 NC mutants (Figure 1c), with the exception of Δ2-NCp9, which did not affect melting force upon DNA binding.
[d]Ty3 NC mutant Δ2-NCp9 did not appreciably bind DNA on the first stretch–release cycle (Figure 5a) over two orders of magnitude in protein concentration (1–100 nM). At protein concentrations 50 nM and above, binding that altered hysteresis was achieved only upon the third stretch–release cycle (Figure 2b and c) and this effect did not change significantly at higher concentrations. Therefore data from the third force-extension curve is included in Table 1 to provide a comparison of melting force and hysteresis ratio for this mutant.

Cationic tail deletions destroy the rapid kinetics and duplex stabilization of wild type Ty3 NC

The N-terminal tail of mutant Δ1-NCp9 is missing the first 16 residues, compromising the protein's DNA binding affinity by ~5-fold (Table 1). K_d was estimated as 12 (±3) nM using a simple DNA binding isotherm to approximate the decrease in melting force as a function of protein concentration [Equations (3) and (5)]. In striking contrast to wild type Ty3 NC, force-extension curves in the presence of 20 nM Δ1-NCp9 exhibit significant hysteresis (Figure 4a). The hysteresis ratio for Δ1-NCp9 is 0.68 (±0.1), relative to 0.26 (±0.03) for wild type (Table 1). This indicates that the mutant does not dissociate from ssDNA on the timescale of the experiment, leading to the loss of rapid kinetics. Subsequent stretches reflect incomplete protein dissociation (Figure 4b). Δ1-NCp9 induces less aggregation than wild type Ty3 NC, and its effect on the force-induced melting transition width is nominal. Furthermore, the melting force is 59.4 (±0.7) pN in the presence of the mutant, which is significantly lower than the 68.0 (±1.0) pN melting force in the presence of wild type Ty3 NC (Table 1). This result reflects weak duplex destabilization by Δ1-NCp9.

To identify the DNA binding activity of the zinc finger, we examined the Δ2-NCp9 mutant. Deletion of the N-terminal tail leaves a mutant that primarily consists of the zinc finger, and has DNA binding affinity more than

an order of magnitude smaller than that of wild type. Although Δ2-NCp9 does not affect DNA melting force or transition width to allow for an estimate of K_d, protein concentrations smaller than 50 nM reflected no DNA binding. In the presence of 50 nM protein, the first force-extension curve lies almost exactly on top of the DNA only curve, showing a negligible effect (Figure 5a). The second stretch–release cycle exhibits some hysteresis, indicating slow kinetics (Figure 5b), which increases further in the third force-extension curve (Figure 5c). The Δ2-NCp9 mutant is the only Ty3 NC protein for which successive stretch-release cycles show additional protein binding. The third stretch-release cycle is required to observe an effect comparable to the other proteins presented in Table 1 because Δ2-NCp9 binds DNA with significantly lower affinity than wild type or the other mutants. Furthermore, these effects remain the same within error, even up to 100 nM protein concentration. Therefore the hysteresis ratio of 0.49 (±0.1) is reported in Table 1 for comparison. There is less hysteresis than that for Δ1-NCp9 (Table 1), which is consistent with the likelihood that Δ2-NCp9 has higher charge density.

DNA stretching curves in the presence of mutants dominated by the zinc finger imply that the zinc finger of Ty3 NC contributes to nucleic acid aggregation and duplex destabilization. Both mutants with a compromised N-terminal tail also lack the rapid kinetics of wild type Ty3 NC, demonstrating that the cationic tail is required to promote nucleic acid annealing.

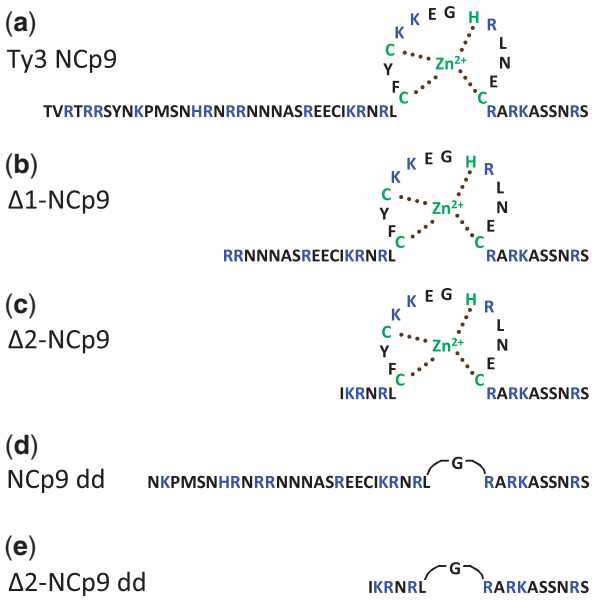

Figure 3. Structure of (**a**) wild-type Ty3 NC, (**b**) Δ1-NCp9, (**c**) Δ2-NCp9, (**d**) NCp9 dd and (**e**) Δ2-NCp9 dd. Basic residues shown in blue and zinc coordinating residues shown in green.

Zinc-finger deletions destroy rapid kinetics and induce strong duplex stabilization

To identify the DNA binding properties of the N-terminal tail, we examined two mutants that lack the zinc finger entirely, NCp9 dd and Δ2-NCp9 dd. The zinc-finger deletion in mutant NCp9 dd leaves the cationic tail nearly intact, resulting in DNA binding affinity most similar to that of wild-type Ty3 NC, with K_d estimated as 3 (±2) nM (Table 1). The initial force-extension curve in the presence of 3 nM NCp9 dd exhibits a significantly higher melting force of 71.5 (±0.4) pN, which reflects strong dsDNA stabilization (Figure 6a, Table 1). The second stretch exactly follows the previous release curve, indicating that the mutant binds ssDNA irreversibly on the timescale of the experiment (Figure 6b). NCp9 dd also induces strong DNA aggregation, indicated by an increase in the DNA stretching force to 10–20 pN below the DNA contour length, an effect that dominates subsequent stretches.

The Δ2-NCp9 dd mutant is Δ2-NCp9 without the zinc finger. Deletion of both the N-terminal tail and the zinc finger results in a small protein likely to have a high charge density, with DNA binding affinity an order of magnitude smaller than that of wild type Ty3 NC. A simple DNA binding isotherm [Equations (3) and (5)] fit to change in melting force ΔF_m as a function of protein concentration c

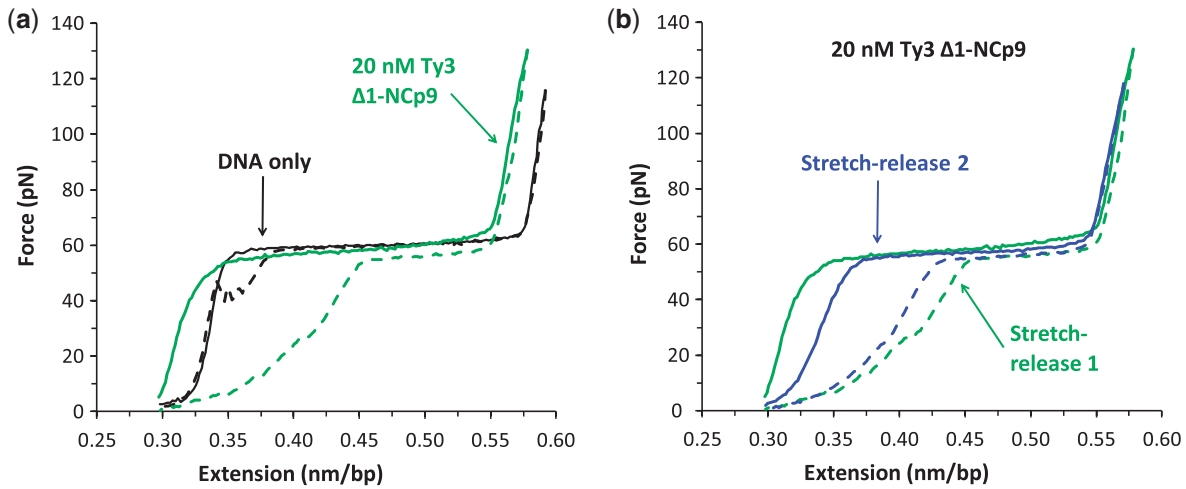

Figure 4. Typical force-extension (solid) and release (dashed) curves of (**a**) DNA only (black) and in the presence of (**a** and **b**) 20 nM Ty3 NC mutant Δ1-NCp9. First stretch–release curve shown in green (**a** and **b**), and second stretch–release curve shown in blue (**b**).

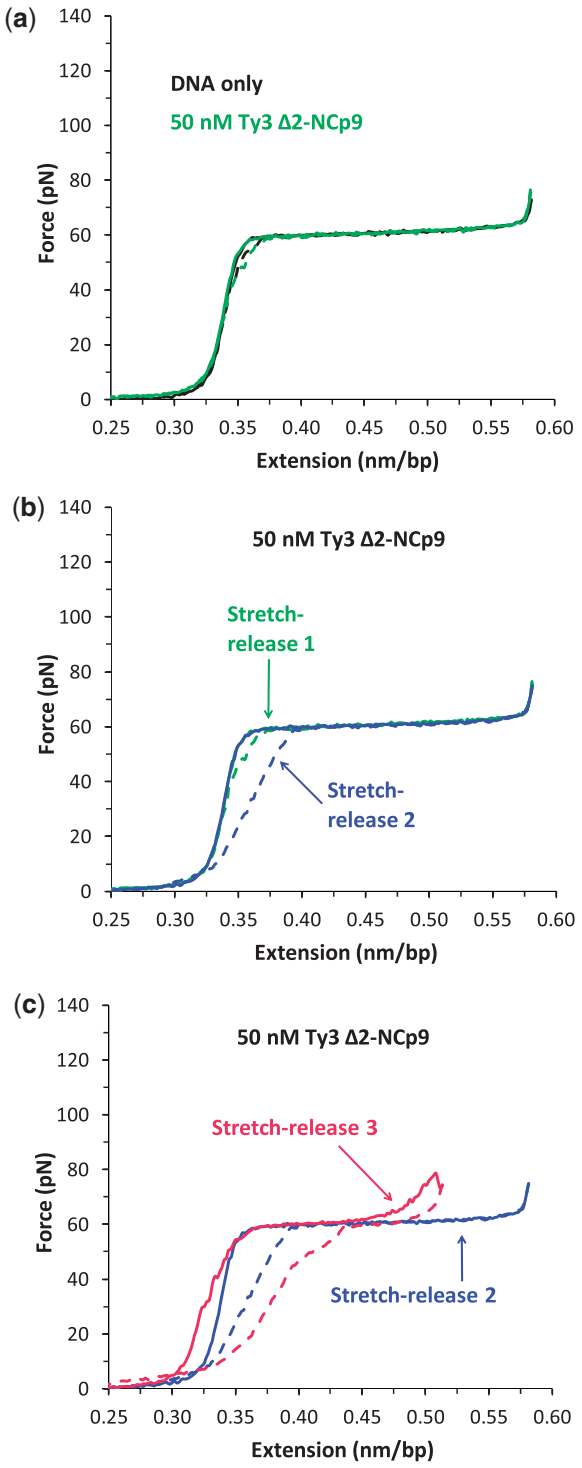

Figure 5. Typical force-extension (solid) and release (dashed) curves of (**a**) DNA only (black) and in the presence of (**a**–**c**) 50 nM Ty3 NC mutant Δ2-NCp9. First stretch–release curve shown in green (**a** and **b**), second stretch–release curve shown in blue (**b** and **c**), and third stretch-release curve shown in red (**c**).

yields K_d = 20 (±1) nM with saturated change in melting force ΔF_m^{sat} = 16 (±0.5) pN. The force-extension curve in the presence of 13 nM Δ2-NCp9 dd exhibits duplex stabilization, DNA aggregation (10 pN force is required to stretch DNA below the dsDNA contour length), and large hysteresis (Figure 7a). In contrast with Δ2-NCp9 (Figure 6b), subsequent stretch–release cycles reflect incomplete protein dissociation (Figure 7b). DNA force-extension curves in the presence of mutants without the zinc finger indicate that the cationic tail of Ty3 NC contributes strongly to nucleic acid aggregation, and is largely responsible for dsDNA stabilization. Both mutants also induce strong hysteresis, and this inhibition of DNA annealing abolishes rapid kinetics, which requires both the zinc finger and the N-terminal tail of Ty3 NC.

DISCUSSION

Recent studies of nucleic acid chaperone proteins from retroviruses and retrotransposons have shown that aggregation, duplex destabilization and rapid kinetics are often characteristic of their chaperone activity. Single-molecule stretching experiments have also shown that an increase in the slope of the DNA melting plateau is positively correlated with nucleic acid chaperone activity (34,45,46). Force-extension curves in the presence of wild type Ty3 NC exhibit this increase in transition width, and demonstrate that aggregation of both single- and double-stranded nucleic acids is a key component of its chaperone activity. The increase in the force at extensions below the DNA contour length has also been observed with retroviral NCs such as HIV-1 NC and RSV NC (34). In addition, the decrease in ssDNA contour length has been reported for ORF1p, the nucleic acid chaperone from the LINE-1 retrotransposon (45,46). DNA stretching curves in the presence of Ty3 NC also exhibit very little hysteresis, indicating that the protein dissociates quickly from ssDNA, allowing the strands to anneal, and rapid kinetics is a key component of its nucleic acid chaperone activity. These results contrast significantly with those obtained in the presence of MuLV NC, the only other single zinc finger NC protein studied by DNA stretching (34). DNA stretching curves in the presence of MuLV NC exhibit significant hysteresis and much less aggregation. Although the previous results suggested that two zinc fingers may be required for rapid DNA interaction kinetics, this study shows that rapid kinetics can be observed even with a single zinc finger NC protein.

We examined several mutants to investigate the relationship between the structure of Ty3 NC and its nucleic acid chaperone function. Stretching curves in the presence of Δ1-NCp9, which has a partially deleted N-terminal tail, indicate that the zinc finger contributes to weak duplex destabilization, an effect that balances the strong duplex stabilization of the cationic tail. Although the distribution of charge depends upon the structure of Ty3 NC mutants bound to nucleic acids, which is unavailable, both mutants without the zinc finger must have high charge density. Both of these mutants, NCp9 dd and Δ2-NCp9 dd, significantly increase the melting force, to 71.5 (±0.4) pN and 67.1 (±0.5) pN, respectively (Table 1). The N-terminal tail mutants demonstrate that the cationic residues on the N-terminus stabilize dsDNA, while the zinc finger destabilizes dsDNA. When the cationic region is deleted, DNA

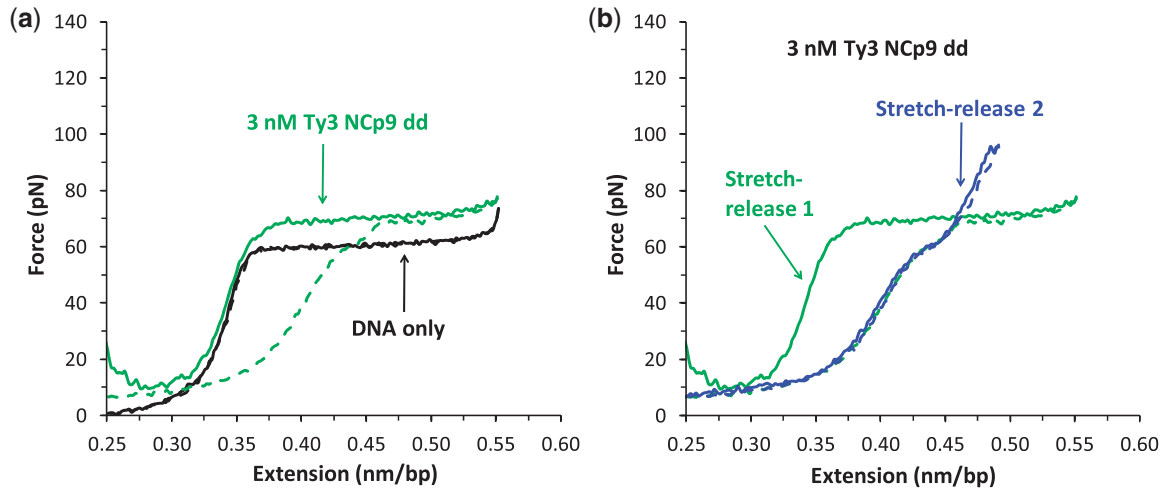

Figure 6. Typical force-extension (solid) and release (dashed) curves of (**a**) DNA only (black) and in the presence of (**a** and **b**) 3 nM Ty3 NC mutant NCp9 dd. First stretch–release curve shown in green (**a** and **b**), and second stretch–release curve shown in blue (**b**).

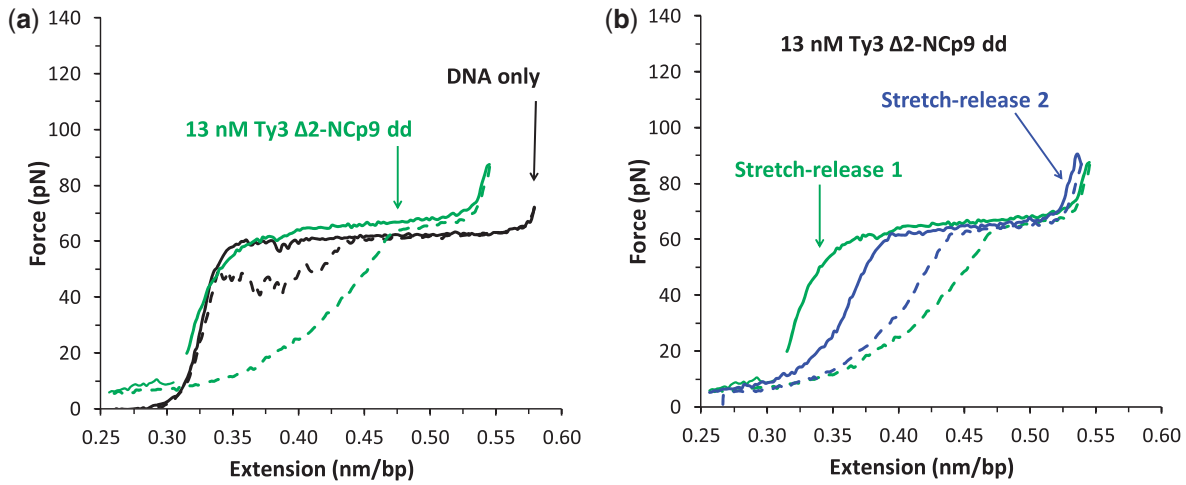

Figure 7. Typical force-extension (solid) and release (dashed) curves of (**a**) DNA only (black) and in the presence of (**a** and **b**) 13 nM Ty3 NC mutant Δ2-NCp9 dd. First stretch–release curve shown in green (**a** and **b**), and second stretch–release curve shown in blue (**b**).

destabilization by the zinc finger, which contains two aromatic residues that could potentially stack with ssDNA, dominates the overall thermodynamics of the DNA-protein interaction. When the zinc finger is deleted, the protein strongly stabilizes dsDNA.

While duplex stabilization such as that observed for wild type Ty3 NC is likely to be somewhat detrimental to nucleic acid chaperone function, this property is a consequence of the high charge density of the protein. A high charge density is required for efficient nucleic acid aggregation, which is a primary requirement for nucleic acid chaperone activity. Thus, wild type Ty3 NC balances aggregation by its cationic residues and duplex destabilization by its zinc finger, resulting in overall mild duplex stabilization in conjunction with strong aggregation. Although all four mutants aggregated both dsDNA and ssDNA, the effect was strongest with zinc finger deletion. Mutant Δ2-NCp9, which is primarily the zinc finger, showed the least DNA aggregation as well as the weakest DNA binding. Even at 100 nM protein concentration, nearly 4-fold higher than saturated binding of wild type Ty3 NC, multiple stretch–release cycles were required to see an effect. In contrast, the zinc-finger deletion of NCp9 dd is likely to result in higher charge density, preserving DNA binding affinity. These properties are consistent with the ability of NCp9 dd to facilitate tRNA primer annealing, promote genomic RNA dimerization, and initiate cDNA synthesis *in vitro* because these processes do not require significant duplex destabilization to proceed (47).

The rapid kinetics of wild type Ty3 NC is highly sensitive to protein charge and structure, and deletion of either the zinc finger or the cationic tail forms mutants that induce hysteresis, which demonstrates that rapid kinetics requires both the zinc finger and the N-terminal tail. In fact, only wild type Ty3 NC has rapid kinetics, quickly

switching between binding single- and double-stranded nucleic acids. The weak duplex destabilization effect of the zinc finger and the strong duplex stabilization effect of the cationic tail work in concert to transiently destabilize secondary structures in the 5′- and 3′-UTR and facilitate annealing of the DNA:RNA duplex during reverse transcription (2,51). Thus, a single zinc finger is sufficient for a highly charged NC protein to exhibit rapid DNA–protein interaction kinetics. Although the high charge density of the protein strongly stabilizes dsDNA, the single zinc finger also destabilizes DNA, and the resulting balance of these two effects makes Ty3 NC a very strong nucleic acid aggregation protein with only weak duplex stabilization. If Ty3 required significant duplex destabilization to facilitate minus strand transfer, as is the case for HIV-1, a second zinc finger would likely be needed. In fact, HIV-1 NC probably evolved a second zinc finger for this purpose, in conjunction with the requirements of other critical HIV-1 replication processes, due to the high thermodynamic stability of its TAR hairpins. This suggests that nucleic acid chaperone activity of NC proteins is specifically tuned to secondary structures in the UTR regions of their genomic RNA. The nucleic acid chaperone properties are then optimized to maximize DNA aggregation while facilitating only the necessary secondary structure rearrangements required for reverse transcriptase to synthesize a complete functional viral DNA.

ACKNOWLEDGEMENTS

The authors would like to thank Micah McCauley and Ioulia Rouzina for insightful discussions. D. Ficheux (IBCP du CNRS, Lyon, France) is gratefully acknowledged for the synthesis of Ty3 NC and its derivatives.

FUNDING

National Institutes of Health (R01GM72462); National Science Foundation (MCB-0744456); Institut National de la Santé et de la Recherche Médicale; Agence Nationale de Recherches sur le Sida et les Hépatitis Virales; National Science Foundation Integrative Graduate Education and Research Traineeship Program (DGE-0504331 to K.C.). Funding for open access charge: National Institutes of Health.

Conflict of interest statement. None declared.

REFERENCES

1. Darlix,J.L., Lapadattapolsky,M., Derocquigny,H. and Roques,B.P. (1995) First glimpses at structure-function relationships of the nucleocapsid protein of retroviruses. *J. Mol. Biol.*, **254**, 523–537.
2. Levin,J.G., Guo,J.H., 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.
3. Rein,A., Henderson,L.E. and Levin,J.G. (1998) Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication. *Trends Biochem. Sci.*, **23**, 297–301.
4. Allain,B., Lapadat-Tapolsky,M., Berlioz,C. and Darlix,J.L. (1994) Transactivation of the minus-strand DNA transfer by nucleocapsid protein during reverse transcription of the retroviral genome. *EMBO J.*, **13**, 973–981.
5. Barat,C., Lullien,V., Schatz,O., Keith,G., Nugeyre,M.T., Gruninger-Leitch,F., Barre-Sinoussi,F., LeGrice,S.F. and Darlix,J.L. (1989) HIV-1 reverse transcriptase specifically interacts with the anticodon domain of its cognate primer tRNA. *EMBO J.*, **8**, 3279–3285.
6. Darlix,J.L., Vincent,A., Gabus,C., de Rocquigny,H. and Roques,B. (1993) Trans-activation of the 5′ to 3′ viral DNA strand transfer by nucleocapsid protein during reverse transcription of HIV1 RNA. *C. R. Acad. Sci., Ser. III*, **316**, 763–771.
7. Prats,A.C., Sarih,L., Gabus,C., Litvak,S., Keith,G. and Darlix,J.L. (1988) Small finger protein of avian and murine retroviruses has nucleic acid annealing activity and positions the replication primer tRNA onto genomic RNA. *EMBO J.*, **7**, 1777–1783.
8. Song,M., Balakrishnan,M., Gorelick,R.J. and Bambara,R.A. (2009) A succession of mechanisms stimulate efficient reconstituted HIV-1 minus strand strong stop DNA transfer. *Biochemistry*, **48**, 1810–1819.
9. Auxilien,S., Keith,G., Le Grice,S.F. and Darlix,J.L. (1999) Role of post-transcriptional modifications of primer tRNALys,3 in the fidelity and efficacy of plus strand DNA transfer during HIV-1 reverse transcription. *J. Biol. Chem.*, **274**, 4412–4420.
10. Peliska,J.A., Balasubramanian,S., Giedroc,D.P. and Benkovic,S.J. (1994) Recombinant HIV-1 nucleocapsid protein accelerates HIV-1 reverse-transcriptase catalyzed DNA strand transfer reactions and modulates RNase-H activity. *Biochemistry*, **33**, 13817–13823.
11. Rodriguez-Rodriguez,L., Tsuchihashi,Z., Fuentes,G.M., Bambara,R.A. and Fay,P.J. (1995) Influence of human immunodeficiency virus nucleocapsid protein on synthesis and strand transfer by the reverse transcriptase in vitro. *J. Biol. Chem.*, **270**, 15005–15011.
12. 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.
13. Berg,J. (1986) Potential metal-binding domains in nucleic acid binding proteins. *Science*, **232**, 485–487.
14. Covey,S.N. (1986) Amino acid sequence homology in gag region of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus. *Nucleic Acids Res.*, **14**, 623–633.
15. Green,L.M. and Berg,J.M. (1990) Retroviral nucleocapsid protein-metal ion interactions: folding and sequence variants. *Proc. Natl Acad. Sci. USA*, **87**, 6403–6407.
16. Henderson,L.E., Copeland,T.D., Sowder,R.C., Smythers,G.W. and Oroszlan,S. (1981) Primary structure of the low molecular weight nucleic acid-binding proteins of murine leukemia viruses. *J. Biol. Chem.*, **256**, 8400–8406.
17. Darlix,J.-L., Godet,J., Ivanyi-Nagy,R., Fossé,P., Mauffret,O. and Mély,Y. (2011) Flexible Nature and Specific Functions of the HIV-1 Nucleocapsid Protein. *J. Mol. Biol.*, **410**, 565–581.
18. De Rocquigny,H., Gabus,C., Vincent,A., Fournie-Zaluski,M.C., Roques,B. and Darlix,J.L. (1992) Viral RNA annealing activities of human immunodeficiency virus type 1 nucleocapsid protein require only peptide domains outside the zinc fingers. *Proc. Natl Acad. Sci. USA*, **89**, 6472–6476.
19. Li,X., Quan,Y., Arts,E.J., Li,Z., Preston,B.D., de Rocquigny,H., Roques,B.P., Darlix,J.L., Kleiman,L., Parniak,M.A. *et al.* (1996) Human immunodeficiency virus Type 1 nucleocapsid protein (NCp7) directs specific initiation of minus-strand DNA synthesis primed by human tRNA(Lys3) in vitro: studies of viral RNA molecules mutated in regions that flank the primer binding site. *J. Virol.*, **70**, 4996–5004.
20. Cimarelli,A. and Darlix,J.L. (2002) Assembling the human immunodeficiency virus type 1. *Cell. Mol. Life Sci.*, **59**, 1166–1184.

21. Darlix,J.L., Gabus,C., Nugeyre,M.T., Clavel,F. and Barre-Sinoussi,F. (1990) Cis elements and trans-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. *J. Biol. Chem.*, **216**, 689–699.
22. Feng,Y.X., Copeland,T.D., Henderson,L.E., Gorelick,R.J., Bosche,W.J., Levin,J.G. and Rein,A. (1996) HIV-1 nucleocapsid protein induces "maturation" of dimeric retroviral RNA in vitro. *Proc. Natl. Acad Sci. USA*, **93**, 7577–7581.
23. Ottmann,M., Gabus,C. and Darlix,J.L. (1995) The central globular domain of the nucleocapsid protein of human immunodeficiency virus type 1 is critical for virion structure and infectivity. *J. Virol.*, **69**, 1778–1784.
24. Sakaguchi,K., Zambrano,N., Baldwin,E.T., Shapiro,B.A., Erickson,J.W., Omichinski,J.G., Clore,G.M., Gronenborn,A.M. and Appella,E. (1993) Identification of a binding site for the human immunodeficiency virus type 1 nucleocapsid protein. *Proc. Natl Acad. Sci. USA*, **90**, 5219–5223.
25. 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.
26. Cruceanu,M., Gorelick,R.J., Musier-Forsyth,K., Rouzina,I. and Williams,M.C. (2006) Rapid kinetics of protein-nucleic acid interaction is a major component of HIV-1 nucleocapsid protein's nucleic acid chaperone function. *J. Mol. Biol.*, **363**, 867–877.
27. Cruceanu,M., Urbaneja,M.A., Hixson,C.V., Johnson,D.G., Datta,S.A., Fivash,M.J., Stephen,A.G., Fisher,R.J., Gorelick,R.J., Casas-Finet,J.R. *et al.* (2006) Nucleic acid binding and chaperone properties of HIV-1 Gag and nucleocapsid proteins. *Nucleic Acids Res.*, **34**, 593–605.
28. 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.
29. 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.
30. Wu,H., Rouzina,I. and Williams,M.C. (2010) Single-molecule stretching studies of RNA chaperones. *RNA Biol.*, **7**, 712–723.
31. Amarasinghe,G.K., De Guzman,R.N., 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.
32. 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.
33. Qualley,D.F., Stewart-Maynard,K.M., Wang,F., Mitra,M., Gorelick,R.J., Rouzina,I., Williams,M.C. and Musier-Forsyth,K. (2010) C-terminal domain modulates the nucleic acid chaperone activity of human T-cell leukemia virus type 1 nucleocapsid protein via an electrostatic mechanism. *J. Biol. Chem.*, **285**, 295–307.
34. Stewart-Maynard,K.M., Cruceanu,M., Wang,F., Vo,M.N., Gorelick,R.J., Williams,M.C., Rouzina,I. and Musier-Forsyth,K. (2008) Retroviral nucleocapsid proteins display nonequivalent levels of nucleic acid chaperone activity. *J. Virol.*, **82**, 10129–10142.
35. Chaurasiya,K.R., Paramanathan,T., McCauley,M.J. and Williams,M.C. (2010) Biophysical characterization of DNA binding from single molecule force measurements. *Phys. Life Rev.*, **7**, 299–341.
36. Housset,V., De Rocquigny,H., Roques,B.P. and Darlix,J.L. (1993) Basic amino acids flanking the zinc finger of Moloney murine leukemia virus nucleocapsid protein NCp10 are critical for virus infectivity. *J. Virol.*, **67**, 2537–2545.
37. Cruceanu,M., Stephen,A.G., Beuning,P.J., Gorelick,R.J., Fisher,R.J. and Williams,M.C. (2006) Single DNA molecule stretching measures the activity of chemicals that target the HIV-1 nucleocapsid protein. *Anal. Biochem.*, **358**, 159–170.
38. Darlix,J., Garrido,J.L., Morellet,N., Mély,Y. and Rocquigny,H.D. (2007) Properties, Functions, and Drug Targeting of the Multifunctional Nucleocapsid Protein of the Human Immunodeficiency Virus. In KuanTeh,J. (ed.), *Advances in Pharmacology*, Vol. 55. Academic Press, San Diego, CA, pp. 299–346.
39. Mely,Y., Rocquigny,H.D., Shvadchak,V., Avilov,S., Dong,C.Z., Dietrich,U. and Darlix,J.-L. (2008) Targeting the Viral Nucleocapsid Protein in Anti-HIV-1 Therapy. *Mini Rev. Med. Chem.*, **8**, 24–35.
40. Kirchner,J. and Sandmeyer,S. (1993) Proteolytic processing of T33 proteins is required for transposition. *J. Virol.*, **67**, 19–28.
41. Gabus,C., Ficheux,D., Rau,M., Keith,G., Sandmeyer,S. and Darlix,J.L. (1998) The yeast Ty3 retrotransposon contains a 5′-3′ bipartite primer-binding site and encodes nucleocapsid protein NCp9 functionally homologous to HIV-1 NCp7. *EMBO J.*, **17**, 4873–4880.
42. Sandmeyer,S.B. and Clemens,K.A. (2010) Function of a retrotransposon nucleocapsid protein. *RNA Biol.*, **7**, 642–654.
43. McCauley,M.J. and Williams,M.C. (2009) Review: optical tweezers experiments resolve distinct modes of DNA-protein binding. *Biopolymers*, **91**, 265–282.
44. Shokri,L., McCauley,M.J., Rouzina,L. and Williams,M.C. (2008) DNA overstretching in the presence of glyoxal: structural evidence of force-induced DNA melting. *Biophys J.*, **95**, 1248–1255.
45. Evans,J.D., Peddigari,S., Chaurasiya,K.R., Williams,M.C. and Martin,S.L. (2011) Paired mutations abolish and restore the balanced annealing and melting activities of ORF1p that are required for LINE-1 retrotransposition. *Nucleic Acids Res.*, **39**, 5611–5621.
46. Martin,S.L., Bushman,D., Wang,F., Li,P.W.L., Walker,A., Cummiskey,J., Branciforte,D. and Williams,M.C. (2008) A single amino acid substitution in ORF1 dramatically decreases L1 retrotransposition and provides insight into nucleic acid chaperone activity. *Nucleic Acids Res.*, **36**, 5845–5854.
47. Cristofari,G., Gabus,C., Ficheux,D., Bona,M., Le Grice,S.F.J. and Darlix,J.L. (1999) Characterization of active reverse transcriptase and nucleoprotein complexes of the yeast retrotransposon Ty3 in vitro. *J. Biol. Chem.*, **274**, 36643–36648.
48. Shokri,L., McCauley,M.J., Rouzina,I. and Williams,M.C. (2008) DNA overstretching in the presence of glyoxal: structural evidence of force-induced DNA melting. *Biophys J.*, **95**, 1248–1255.
49. van Mameren,J., Gross,P., Farge,G., Hooijman,P., Modesti,M., Falkenberg,M., Wuite,G.J.L. and Peterman,E.J.G. (2009) Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging. *Proc. Natl Acad. Sci. USA*, **106**, 18231–18236.
50. Williams,M.C., Rouzina,I. and Bloomfield,V.A. (2002) Thermodynamics of DNA interactions from single molecule stretching experiments. *Accounts Chem. Res.*, **35**, 159–166.
51. Cristofari,G. and Darlix,J.L. (2002) The ubiquitous nature of RNA chaperone proteins. *Prog. Nucleic Acid Res. Mol. Biol.*, **72**, 223–268.