RNA Biology 7:6, 712-723; November/December 2010; © 2010 Landes Bioscience

Single-molecule stretching studies of RNA chaperones

Hao Wu,[1] Ioulia Rouzina[2] and Mark C. Williams[1,3,*]

[1]Department of Physics; Northeastern University; Boston, MA; [2]Department of Biochemistry, Molecular Biology and Biophysics; University of Minnesota; Minneapolis, MN USA; [3]Center for Interdisciplinary Research on Complex Systems; Northeastern University; Boston, MA USA

Key words: force spectroscopy, SSB protein, LINE-1, ORF1p, DNA condensation, nucleocapsid protein, NC

RNA chaperone proteins play significant roles in diverse biological contexts. The most widely studied RNA chaperones are the retroviral nucleocapsid proteins (NC), also referred to as nucleic acid (NA) chaperones. Surprisingly, the biophysical properties of the NC proteins vary significantly for different viruses, and it appears that HIV-1 NC has optimal NA chaperone activity. In this review we discuss the physical nature of the NA chaperone activity of NC. We conclude that the optimal NA chaperone must saturate NA binding, leading to strong NA aggregation and slight destabilization of all NA duplexes. Finally, rapid kinetics of the chaperone protein interaction with NA is another primary component of its NA chaperone activity. We discuss these characteristics of HIV-1 NC and compare them with those of other NA binding proteins and ligands that exhibit only some characteristics of NA chaperone activity, as studied by single molecule DNA stretching.

Introduction

Nucleic acid (NA) chaperones are proteins or small molecules that facilitate the rearrangement of nucleic acids structures into their lowest energy states, by optimizing the number of base pairs and tertiary interactions.[1,2] Several different types of NA chaperones are known, as described in this special issue. Some NA chaperone proteins facilitate the native NA fold by specifically binding to and stabilizing this fold, while others work by preventing specific alternative NA structures. In this chapter, we discuss a different type of NA chaperone, which acts sequence non-specifically on any NA, leading to a rapid conversion to minimal free energy conformations. This type of NA chaperone appears indispensable for facilitation of reverse transcription (RTion) in retroviruses and retrotransposons. The process of copying the retroviral genomic RNA (gRNA) into proviral dsDNA involves continuous remodeling of NAs, including multiple complementary NA sequence annealing steps, many of which must be converted from a highly structured state. Interestingly, both retroviruses and retrotransposons encode their own NA chaperones, thereby highlighting the critical importance of these proteins in the retroviral life cycle. In the case of retroviruses, these NA chaperones constitute the RNA-binding domain of the Gag polyprotein in the immature capsid. In the course of virus maturation, the nucleocapsid (NC) domain is cleaved from Gag, and serves as a major facilitator of RTion. In contrast to more specific NA chaperone proteins, NC does not use ATP or another energy source. However, stoichiometric binding of NC to the entire gRNA is necessary for its chaperone function. It is this type of saturated protein binding to polymeric NA that is most amenable for studies using single molecule (SM) stretching of long dsDNA molecules, as is described below.

*Correspondence to: Mark C. Williams; Email: mark@neu.edu
Submitted: 08/03/10; Revised: 09/15/10; Accepted: 09/16/10
Previously published online
www.landesbioscience.com/journals/rnabiology/article/13776
DOI: 10.4161/rna.7.6.13776

DNA Force-Induced Melting Studies Using Optical Tweezers

Figure 1A is a schematic diagram of a dual beam optical tweezers instrument for stretching single DNA molecules. In these experiments, two streptavidin-coated polystyrene beads are attached to both sides of one double-stranded bacteriophage λ genomic DNA molecule (48,500 bp) by interacting with biotin molecules at the 5' end of each DNA strand. Two laser beams are overlapped to form an optical trap to trap one bead, and the other bead is attached to a glass micropipette. By moving the glass micropipette, the DNA is stretched. The force required to stretch the DNA a certain distance is obtained, resulting in a measurement of the force-extension profile of the DNA molecule, shown in **Figure 1B**.

In the stretching process, initially a small force (approximately 5pN) is applied to stretch DNA to its B-form contour length. After DNA is straightened, a rapid force increase is required to stretch the duplex DNA structure. In this regime, the curve is consistent with theoretical dsDNA stretching curve from the Worm-Like Chain (WLC) model,[3] a model that accurately represents the elasticity of a rigid polymer such as dsDNA. At a force of approximately 60 pN, a phase transition from dsDNA to ssDNA occurs. In this regime, DNA can be stretched to 1.7 times its B-form contour length with very little increase in force, shown as a plateau in **Figure 1B**, and the fraction of the ssDNA increases as the length of the construct increases through the transition. After this transition, dsDNA is transformed into ssDNA, and the stretching curve goes up following the theoretical ssDNA stretching curve from the Freely-Jointed Chain (FJC) model, a model that fits to ssDNA stretching with reasonable but not perfect accuracy.[4]

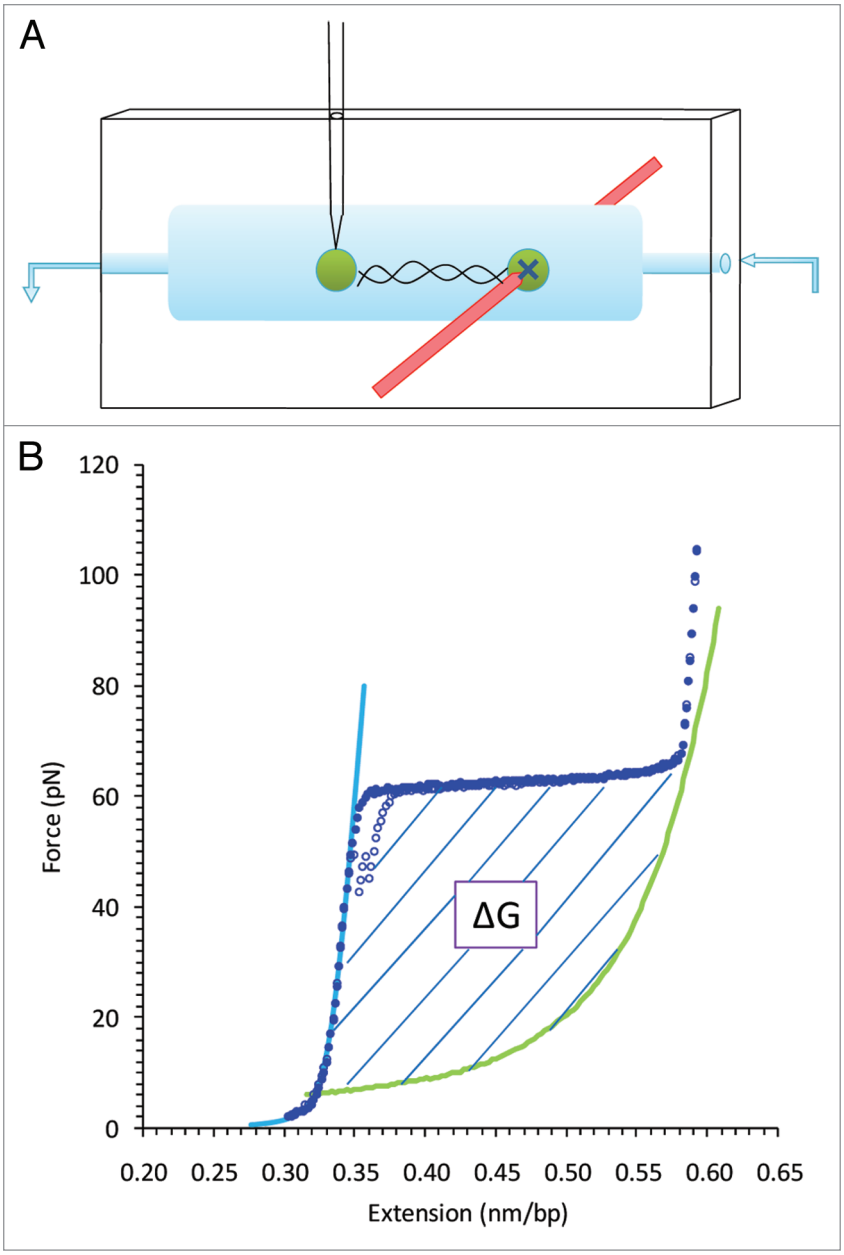

Figure 1. (A) Schematic diagram of a dual beam optical tweezers instrument. (B) DNA force-extension curves. Solid circles represent the stretching curve of naked λ-DNA, while open circles are the relaxation curve of naked DNA. The blue curve is the theoretical force-extension curve of dsDNA, based on the wormlike chain model. The green curve shows the force-extension curve of ssDNA.

As shown in **Figure 1B**, without any proteins or ligands interacting with the DNA (except monovalent salt), this stretching and relaxation process can be considered reversible. When this is the case, the work done by the external force to stretch DNA as a double strand, and then to relax it as a single strand, i.e., the area surrounded by the WLC model, transition plateau and FJC model, labeled by ΔG, is the free energy of DNA melting. Analogous to the dsDNA thermal melting temperature T_m, the overstretching transition midpoint force, F_m, is affected by solution conditions such as high and low pH value, high temperature and low salt.[5-8] In addition, some small proteins or ligands with the function of NA stabilization or destabilization can affect the overstretching force, as expected for DNA melting.[9-12] Therefore, the DNA overstretching transition should properly be considered as a DNA force-induced melting transition.

We also know that DNA is melted by force and the DNA base pairs are exposed during this transition based on several recent experiments using glyoxal, intercalating dyes and fluorescently labeled single stranded binding proteins (SSBs). Glyoxal ($C_2H_2O_2$) is a small molecule which binds irreversibly to exposed guanine bases of DNA with slow kinetics and therefore hinders the reannealing of the DNA. In the experiment, λ-DNA molecules were held at fixed extensions for 30 min in the presence of glyoxal, which is the timescale required for DNA binding. When initially stretched in the absence of glyoxal, the DNA stretching curve showed a contour length equal to that expected for dsDNA, 0.34 nm/bp. Upon relaxation, the force-extension curve was almost identical to that obtained upon stretching. However, after treatment with glyoxal, the force-extension curve upon relaxation exhibited a much longer contour length, as expected for a DNA molecule that is partially ssDNA and partially dsDNA. This disagreement between stretching and relaxation curves is referred to as hysteresis. Such hysteresis is often observed when DNA is unable to reanneal after being melted by force, as observed after glyoxal treatment. What is unusual in the case of glyoxal is that subsequent stretches also show an increased contour length because the DNA is essentially frozen in the melted state due to the fact that guanine bases are covalently modified, which subsequently prevents DNA reannealing. As will be shown later, in the presence of DNA binding proteins, more temporary hysteresis is an indication of slow protein dissociation from ssDNA. Recent single-molecule studies have also directly visualized the nature of the structural transition using a combination of optical tweezers and fluorescence imaging techniques.[13] A DNA molecule is stretched in the presence of YOYO, a fluorescent dye which only intercalates into the paired bases of dsDNA. Therefore, the length of DNA to which YOYO can bind reveals the fraction of dsDNA. This experiment visually illustrates a structural conversion from dsDNA into ssDNA. In addition, the experiments on Human mitochondrial SSB which only binds and wraps relaxed ssDNA and the SSB Replication Protein A, which only binds ssDNA, also provide evidence to support the force-induced melting model. Although the latter study showed that DNA only melted by peeling from the ends, it is likely that in the presence

of binding ligands melting also occurs in the middle of the DNA molecule, as is the case for thermal melting.[13-16]

Single Molecule Studies of Cationic Ligands

The dsDNA molecules of most organisms are very long semi-rigid molecules that in vivo must fit into a very small spaces such as the eukaryotic nucleus. Similarly, the two copies of genomic retroviral RNA must fit into the small volume of the viral capsid, which probably remains intact during much of the RTion process.[17-21] Usually several-hundred-fold linear compression is required for such NA packaging, and therefore there are many energetic barriers to overcome in this process, such as the loss of configurational entropy of the long NA molecule, the tight bending of the stiff double helix and the electrostatic repulsion of the negatively charged NA.[22] Therefore some mechanisms to lower these energetic barriers, decrease persistence length and facilitate aggregation must be relevant in vivo. Indeed, it is known[23] that physiologically relevant small multivalent cations with a net charge of +3 or higher, such as Spermidine^{3+} or Spermine^{4+},[22-32] induce aggregation (or condensation) of dsDNA and RNA.[33,34] In addition, protamines,[35] synthetic multivalent cations (such as CoHex^{3+}), and many unstructured cationic peptides or proteins with high positive charge density also condense DNA and RNA. Interestingly, HIV-1 NC was shown to aggregate both ssRNA and dsDNA[36-41] in a very similar way.

It has previously been shown that aggregation of long heat-melted single-stranded DNA molecules by multivalent cations increased the DNA re-annealing rate by up to 6 orders of magnitude.[42] This is because the decrease in the volume of the bimolecular reaction upon condensation makes it easier for the complementary NA regions to find each other. While there are many ligands and proteins that are able to aggregate NA, not all of them are equally efficient in facilitating annealing. For example, most cationic lipids, NA-binding self-aggregating proteins and proteins that crosslink distant specific regions on dsDNA strongly compact NA. In most of these cases, however, the NA has low mobility within the aggregates, leading to a reduced potential for refolding. In contrast, the NA aggregates produced by multivalent cations are characterized by the highest possible NA density along with high NA and cation mobility. These features are the result of the mechanism of electrostatic aggregation, which occurs because of the shrinking of the overall neutral "ionic crystal" of compact positive charges on the background of the "smeared" negative charge of NA. This type of aggregation occurs only when the multivalent cations are bound to the entire NA structure in a primarily delocalized manner, such that they can freely move to optimize their electrostatic free energy.[25,43-46] There is, therefore, a direct connection between the rapid kinetics of the NA chaperone-NA interaction and NA aggregation that optimally facilitates NA refolding. Just as saturating amounts of multivalent cations are required for electrostatic aggregation,[25,29] saturating HIV-1 NC is needed for optimal chaperone activity.[39]

In single molecule DNA stretching experiments, multivalent cation-induced dsDNA aggregation is observed as a small constant "de-aggregating" force of 1–10 pN required for unwrapping the dsDNA wrapped on itself in a toroid from its maximum compaction to the dsDNA contour length.[47-51] Examples of such low de-aggregating forces observed below the dsDNA contour length are presented in **Figure 2B** for DNA stretching with Poly-lysine[51] and later for DNA stretching with the ORF1p protein (see below).[52] This force is a direct measure of the multivalent cation-induced DNA-DNA attractive energy per unit length,[53] and it depends on the cation's net charge and its density, as well as on the concentration of competing monovalent salt in solution.[48,53] Interestingly, dsDNA de-aggregating forces similar to those measured in the presence of multivalent cations were observed in the presence of HIV-1 NC.[51] However, the de-aggregating force was less constant than in the case of simple cations (CoHex^{3+} or Spermine^{4+}) and showed some ruggedness, most likely, reflecting some non-electrostatic protein-DNA interactions, resulting in lower protein mobility on DNA. dsDNA de-aggregating forces can therefore be considered as a signature of dsDNA aggregation by a particular ligand.

In addition to inducing aggregation of almost any nucleic acids, including DNA or RNA in double-, single- or triple-stranded forms, multivalent cations also cause a significant increase in the flexibility of dsDNA.[47] This effect was discovered by careful analysis of the shape of the dsDNA stretching curves near the dsDNA contour length,[47] from which one can obtain dsDNA persistence length measurements under aggregating solution conditions. It was previously known[24,54] that high concentrations of monovalent cations enhance dsDNA flexibility by screening its charge, leading to saturation of the DNA persistence length at a value of 50 nm.[55] However, multivalent cations reduce the dsDNA persistence length to 25–30 nm and the strength of this effect correlates with the DNA aggregating ability of the multivalent cation,[56] making the aggregation energetically more favorable.

These results suggest that the ability of multivalent cations to aggregate NA is analogous to that of NC proteins, which facilitate NA refolding. However, in contrast to HIV-1 NC, which destabilizes NA duplexes,[57-60] multivalent cations stabilize duplexes by contributing to higher solution ionic strength, thereby providing better screening for the higher charge on dsNA relative to ssNA. This effect leads to an increase in the melting temperature of NA duplexes,[54,61] as well as to an increase in the dsDNA melting force.[62]

Several studies suggest that the duplex destabilizing activity of HIV-1 NC is associated with its two properly folded zinc fingers, which stack their aromatic residues (W and F) with unpaired NA bases, leading to preferential HIV-1 NC binding to ssNA duplexes.[51,57,60,63-66] Therefore, we expect the zinc-less version of HIV-1 NC to lose this ability and to behave more like a multivalent cation in DNA stretching experiments. Indeed, as illustrated by **Figure 2**, the DNA stretching curves in the presence of HIV-1 SSHS NC, a mutant form of NC that is unable to coordinate Zn (**Fig. 2A**), resemble DNA stretching curves in the presence of the unstructured cationic polypeptide poly-lysine (**Fig. 2B**).[57] In both cases the DNA melting force is significantly raised by the presence of the peptide, while the curve's shape remains almost unaffected. In addition, in both cases some DNA aggregation is apparent at extensions below the dsDNA contour length.

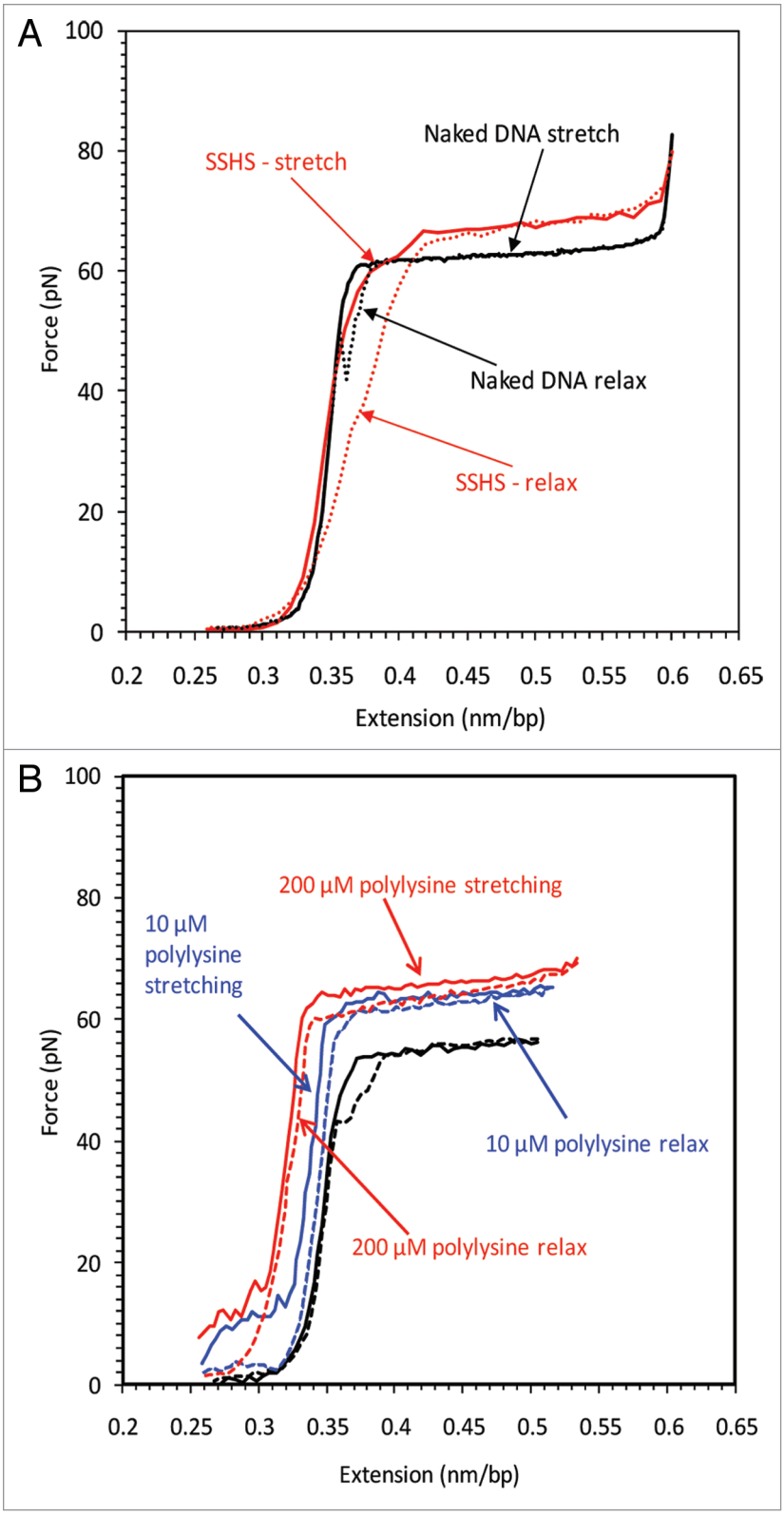

Figure 2. (A) DNA stretching (solid line) and relaxation (dotted line) curves in the absence of protein (black) and in the presence of SSHS NC (red). (B) DNA stretching (solid line) and relaxation (dashed line) curves in the presence of 200 µM polylysine (red) and 10 µM polylysine (blue).

This is in contrast to the DNA stretching curves in the presence of wt HIV-1 NC (see below) that are characterized by a significant lowering of the midpoint force of the DNA melting transition, as well as by an increased transition slope, which likely reflects a decrease in the DNA meting cooperativity and some DNA sequence-specific binding.[57] These results are in agreement with in vitro experiments, which showed that wild type HIV-1 NC was only better than SSHS NC at annealing the highly structured complementary regions of HIV-1 long repeats involving stable TAR hairpins.[67] In contrast, SSHS NC and wild type HIV-1 NC showed similar annealing activity for the much less structured complementary 18-nt primer binding site sequences in HIV-1 plus-strand transfer.[67] The presence of SSHC NC also greatly increases the efficiency of tRNA annealing by facilitating strand attraction.[68] Because tRNA annealing does not require significant tRNA destabilization, the zinc fingers are not needed. In contrast, because HIV-1 minus strand transfer requires destabilization of NA structures containing many base pairs, NC's zinc fingers are required. Finally, the ability of NC proteins to aggregate NA to their maximum density may play another important role in the retroviral life cycle by possibly keeping the mature capsid associated in the cytoplasm of the infected cell throughout much of the RTion process.[17-21]

Single Molecule Studies of Single-Stranded DNA Binding Proteins

In vivo, DNA strands form a helical duplex structure to protect the hydrophobic base pairs from solution. However, during some specific processes such as DNA replication, repair and transcription, the helical duplex DNA structure must be melted, leaving the ssDNA exposed and vulnerable. SSB proteins can protect ssDNA from nucleases, chemical degradation and transient secondary structure formation by binding to exposed ssDNA as it is formed.[14] Therefore, such proteins play important roles in the life cycles of viruses and cells. Different SSB proteins vary greatly in size, complexity and binding mechanism. Among them, bacteriophage T4 gp32 and T7 gp2.5 represent relatively simple model systems that have been well studied using single molecule and other methods.[14]

It is known that one of the simplest and most studied SSB proteins, T4 gp32, has some NA chaperone properties and attempts to use these properties for optimizing in vitro transcription and reverse transcription assays are ongoing. A recent study using gp32,[69] has demonstrated a 2-fold increase in the production of full-length cDNA in in vitro

gRNA reverse transcription as compared to the previous best yields obtained at elevated temperatures. According to the study,[69] the primary reason that gp32 facilitates transcription is by preventing the formation of template ssDNA and RNA secondary structures, allowing for faster polymerization of the complementary strand by the polymerase or reverse transcriptase. This facilitation of the full-length cDNA production by gp32 is significantly less than the analogous effect of HIV-1 NC in in vitro reverse transcription assays.[70] The reason for the relatively poor NA chaperone activity of gp32 becomes clear as we consider the results of single molecule DNA stretching studies of gp32-DNA interactions[71-74] that are briefly described below.

T4 gp32 has a DNA binding site containing an oligonucleotide binding (OB) fold in the core, an N-terminal domain used for highly cooperative binding to ssDNA and a C-terminal domain (CTD) that interacts with other replication proteins. **Figure 3A** shows the force extension curves of DNA in the presence of T4 gp32 and its CTD truncation mutant T4 gp32 *I obtained at a DNA pulling rate of 100 nm/s. In contrast to DNA stretching curves in the presence of mono- or multivalent cations or HIV-1 NC, the DNA stretching curves in the presence of gp32 depend strongly on the DNA pulling rate. Also, in contrast to that observed with HIV-1 NC, the DNA force-extension curves with gp32 appear very different upon DNA stretching through the melting transition relative to the subsequent relaxation of DNA. Indeed, the DNA stretching curves show very little effect of gp32, as if the protein did not have enough time to bind. On the other hand, the DNA relaxation curves indicate that the DNA remains single-stranded and permanently bound by the protein, such that no strand re-annealing after melting can occur. When the DNA stretch-relax cycle in the presence of gp32 is performed at a slower rate, the force-extension curves in the stretch and relax part of cycle become closer to each other, and are expected to eventually converge in a single hysteresis-free equilibrium curve. However, gp32's ssDNA on and off rates are so slow that this equilibrium curve is unobtainable on the time scale of our experiments. This result allows us to estimate the on and off rates, as discussed below. Moreover, this extreme case of a very "slow" protein helps us to understand that the amount of hysteresis in the DNA stretch-relax cycle and the dependence of the curves on the DNA pulling rate are both quantitative signatures of the rates of protein-DNA interaction. These non-equilibrium features of DNA stretching curves in the presence of various proteins allow for semi-quantitative analysis of the

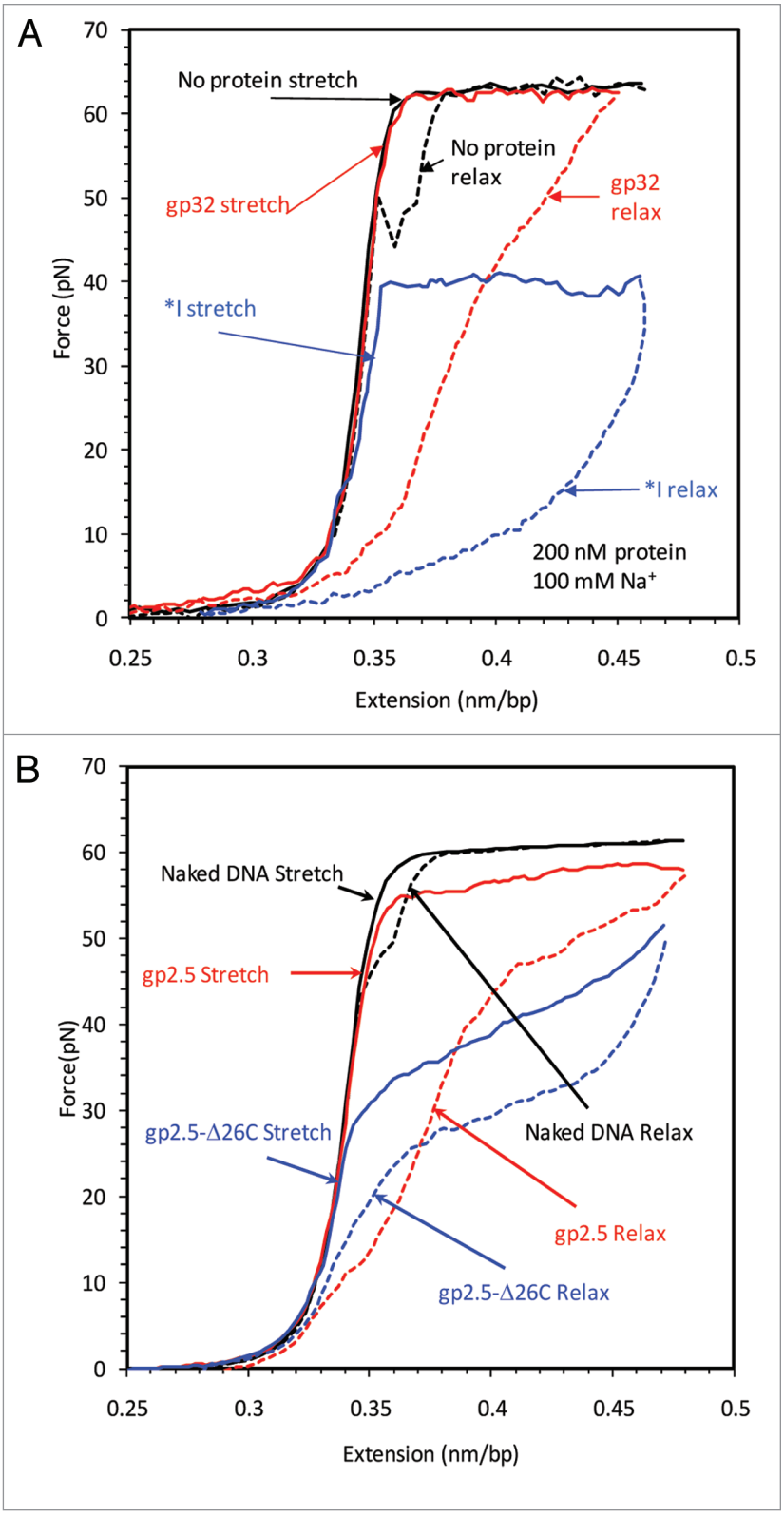

Figure 3. (A) DNA stretching (solid line) and relaxation (dashed line) curves in the absence of protein (black) and in the presence of T4 gp32 (red) and its mutant T4 gp32 *I (blue). (B) DNA stretching (solid line) and relaxation (dashed line) in the absence of protein (black) and in the presence of T7 gp2.5 (red) and its mutant T7 gp2.5-Δ26C (blue).

kinetics of the protein-DNA binding events for any chaperone protein. Interestingly, the kinetics of the protein-NA interaction appears to be a major determinant of a particular protein's NA chaperone activity.

In an extensive series of single molecule DNA stretching studies[71-75] on gp32, we developed a methodology that allows for the complete quantitative characterization of the equilibrium and kinetic properties of a protein's DNA interactions. In particular, fixing the DNA length near the melting transition midpoint and watching relaxation of the force over time allows us to estimate the equilibrium melting force, F_m and the relaxation rates associated with the protein's on and off processes. The slow off times, on the order of several minutes for gp32 dissociation from ssDNA, are due to the cooperative binding of this protein to ssDNA, leading to its preferential dissociation only from the end of the gp32 bound cluster at the dsDNA-ssDNA boundary. Moreover, performing such measurements as a function of protein concentration allowed us to construct the gp32-ssDNA binding isotherm, and to determine its ssDNA binding constant. Furthermore, measuring the DNA strand melting force vs. pulling rate allowed us to estimate the protein's ssDNA on rate under the condition when all of the available ssDNA is already bound by polymerized gp32, while the new ssDNA site for gp32 appeared upon stretching the DNA. The measured on rate dependence on gp32 concentration suggests that this protein finds its new ssDNA site by sliding along the dsDNA. This observation allowed us to determine the gp32 equilibrium binding constant to dsDNA. Finally, all of these characteristics of gp32 were studied as a function of monovalent salt concentration for wt gp32 and its CTD truncation mutant *I, allowing us to construct a detailed model of the electrostatic regulation of gp32-DNA binding by this protein's anionic CTD. This model is in good agreement with the results of all other more conventional studies of this protein, but also provides extensive new information.

The technique described above that was originally developed for the analysis of gp32-DNA interactions was subsequently successfully applied to another SSB protein, T7 gp2.5.[10,76,77] Even a superficial comparison of the DNA stretching and relaxation curves in the presence of gp32 and gp2.5 in **Figure 3A and B**, which were obtained at the same DNA pulling rate of 100 nm/s, leads to the conclusion that gp2.5 exhibits faster DNA interaction kinetics relative to gp32. Interestingly, many features of gp32-DNA interactions are also present for gp2.5. In particular, this SSB protein also has a highly anionic CTD, which significantly slows down this protein's DNA on rate. This can be observed by comparing the DNA stretch/release curves in the presence of wt gp2.5 and its CTD truncation mutant T7 gp2.5-Δ26C shown in **Figure 3B**. Similar to T4 gp32, its C-terminal truncation mutant exhibits a stronger ability to lower the DNA melting force, which can be also explained by its stronger preferential binding to ssDNA. Both C-terminal tails have been shown to weaken ssDNA binding, but they work by different mechanisms. In the case of T4 gp32, the C-terminal tail binds directly to the DNA binding site of the same molecule, inhibiting ssDNA binding. In the case of T7 gp2.5, the C-terminal tail binds to the ssDNA binding site of its dimer partner, facilitating dimerization and inhibiting ssDNA binding. In both cases, removal of the C-terminus exposes the DNA binding site on the protein core and facilitates ssDNA binding by the protein, which is consistent with our single molecule experiment results. In addition, the smaller hysteresis indicates that T7 gp2.5 has a faster off rate than T4 gp32, which is on the order of tens of seconds.[10,76,77]

The results above demonstrate that all four of the tested SSB proteins have the ability to destabilize dsDNA in equilibrium and the mechanism of destabilization can be quantified by DNA stretching. Importantly, all of these SSB proteins appear to have slow ssDNA on and off kinetics (when compared to the best NA chaperone, HIV-1 NC). This property of SSB proteins emerges as a primary reason for their relatively weak NA chaperone activity as characterized by in vitro reverse transcription assays,[69,70] despite their strong duplex destabilizing activity. Moreover, excessive duplex destabilization by some SSB proteins may lead to complete equilibrium melting of the DNA duplex by the protein. This can be detrimental for the NA chaperone function of an SSB protein, which should facilitate the folding of an NA structure into its most stable base-paired conformation. For this to occur, the final base-paired structure must be stable in the presence of the protein.

NA Chaperone Proteins: Retroviral NC Proteins

After the retrovirus infects the host cell, its RNA must be copied into dsDNA through RTion. This process is performed by the enzyme reverse transcriptase (RT) with the participation of NC proteins. RTion contains three obligatory NA annealing steps: tRNA primer annealing, minus-strand transfer and plus-strand transfer. Studies have shown that all these three steps need the participation of NC proteins, especially the minus stand transfer step.[66,78-80] The effect of NC's chaperone activity on these processes is detailed in the review by Levin and Musier-Forsyth in this special issue.

In addition to these obligatory strand-transfer events, each RTion process involves on average 10 random strand transfers[81] in which RT switches between the two viral RNA templates. These events are extremely important for viral survival, as each of the two ~10,000-nucleotide ssRNA genomes often appears to be broken, necessitating strand transfer. In addition, random strand transfers lead to extensive recombination between the two non-identical gRNA molecules,[81-84] facilitating rapid retroviral evolution. During each strand transfer event, the newly synthesized cDNA strand has its template RNA degraded by the RNase-H activity of RT and anneals to the complementary region on the other gRNA molecule, leading to a switch of RT to another template.[85-88] All of these processes are bimolecular in nature, and involve annealing of extensively structured complementary RNA and DNA regions. Therefore, all of them rely heavily on the chaperone activity of NC proteins present at high concentrations in all retroviruses.[57,66,84,89-91] The major role of HIV-1 NC in strand transfer events was also confirmed in a recent in vitro RTion study.[70] The authors showed that it is the efficiency of strand transfers, which occur on average every 300 nt, that greatly increase the

yield of the full-length cDNA production, rather than the RT polymerization rate or processivity. Moreover, it was shown that RNA aggregation by NC is needed for this activity, and that the gRNA within the aggregate is fully accessible to RT, even if RT was added after the RNA aggregation had occurred. All of these features illustrate the major role for reverse transcription of the NA aggregating ability of HIV-1 NC, together with the high mobility of RNA within these dense aggregates.

Interestingly, single-molecule DNA stretching experiments allow for characterization of all three aspects of NC's NA chaperone activity: NA aggregation, duplex destabilization and kinetics of interaction, which we will illustrate using the well studied HIV-1 NC protein as an example. HIV-1 NC consists of 55 amino acids with two nonequivalent CCHC zinc fingers and one positively charged N-terminal tail and each zinc finger contains one aromatic residue. **Figure 4** shows a stretching and relaxation curve of DNA in the presence of a saturating amount (10 nM) of wild type HIV-1 NC. Comparing the NC-free and NC-bound DNA stretching curves in **Figure 4** we notice that this protein significantly lowers the midpoint force of the DNA melting transition. Calculation of the work performed by the force to melt dsDNA provides an estimate of the average base pair stability in the presence of HIV-1 NC. While in the absence of HIV-1 NC each bp stability is on an average ~1.5 kcal/mol/bp, in the presence of HIV-1 NC it is reduced 30% to only ~1 kcal/mol/bp,[92] thereby classifying HIV-1 NC as a moderate duplex destabilizer. This moderate duplex destabilizing ability of HIV-1 NC is central to its optimal chaperone activity. Indeed, while still weaker duplex destabilization would slow down NA refolding by stabilizing misfolded intermediates, stronger duplex destabilization might lead to unfolding of the final annealed NA structure by the protein chaperone. Thus, the modest destabilization of every NA bp by HIV-1 NC is equivalent to the effect of a moderate temperature increase to about ~20°C below the duplex T_m, which is known to lead to the optimal NA annealing rates.[93,94] Another prominent feature of the DNA stretching curves in the presence of saturating HIV-1 NC is the broadening of the DNA melting force plateau. This feature may indicate any of the three following effects: (i) strong HIV-1 NC sequence specificity for ssDNA binding, (ii) decrease in the DNA melting cooperativity, due to HIV-1 NC, (iii) extensive intercalation of aromatic residues of HIV-1 NC into dsDNA at forces above ~30 pN, but below melting force.

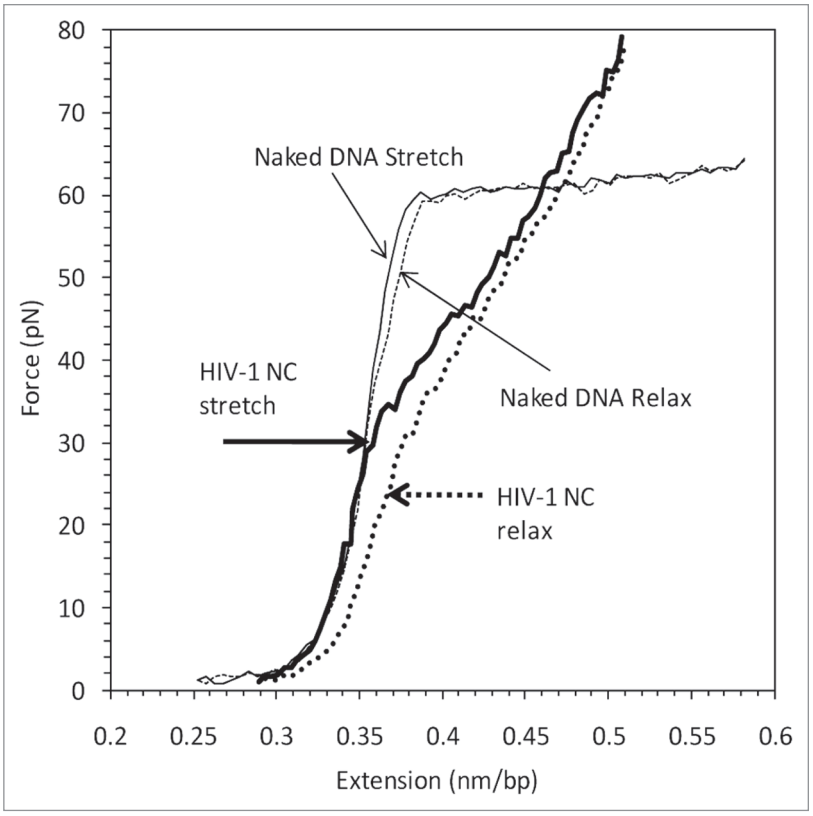

Figure 4. DNA stretching (solid line) and relaxation (dotted line) curves in the absence of protein (thin black line) and in the presence of 10 nM HIV-1 NC (thick black line).

An even more important feature of the data is that in the presence of saturated HIV-1 NC the DNA stretching and relaxation curves closely follow each other (**Fig. 4**), showing that as DNA is relaxed and tends to re-anneal, HIV-1 NC rapidly dissociates from ssDNA and re-binds dsDNA, allowing for the equilibrium state of DNA at every extension and leading to very small hysteresis. In general, small hysteresis in the DNA stretch/relax cycle indicates rapid protein-DNA interaction kinetics. Taken together, single molecule DNA stretching studies of various HIV-1 NC mutants,[51,92] HIV-1 NCp7 precursor forms Gag, NCp15 and NCp9,[41] as well as of NC proteins from other retroviruses[95] suggest that the binding of most of these proteins to DNA result in much larger DNA stretching hysteresis, indicating much slower protein-DNA interaction kinetics. The rapid protein-DNA interaction kinetics, as characterized by small hysteresis in DNA stretching and relaxation experiments, appears to be the physical property of NC proteins that most strongly correlates with the in vivo infectivity of the corresponding mutant viruses.[51] The precise mechanism by which HIV-1 NC maintains its rapid NA interaction kinetics on the molecular level is still a mystery. Indeed, the protein's duplex destabilizing ability requires its preference for unpaired NA bases. This preference, in turn, requires protein-NA interactions other than nonspecific electrostatic binding. Apparently, HIV-1 NC is capable of balancing these two contradictory properties to optimize its NA chaperone function. This balancing act relies on finely tuned stacking of the HIV-1 NC's aromatic residues with the unpaired NA bases, such that even minor alteration of these essential residues or of their positioning within NC's zinc fingers leads to slower protein-NA interaction kinetics and poor chaperone function. More studies are underway to understand the molecular details of this balancing act.

Therefore, single molecule DNA stretching studies provide a simple and straightforward approach for screening various proteins and ligands for their NA chaperone activity. It can also be used for testing drugs that may interfere with the chaperone

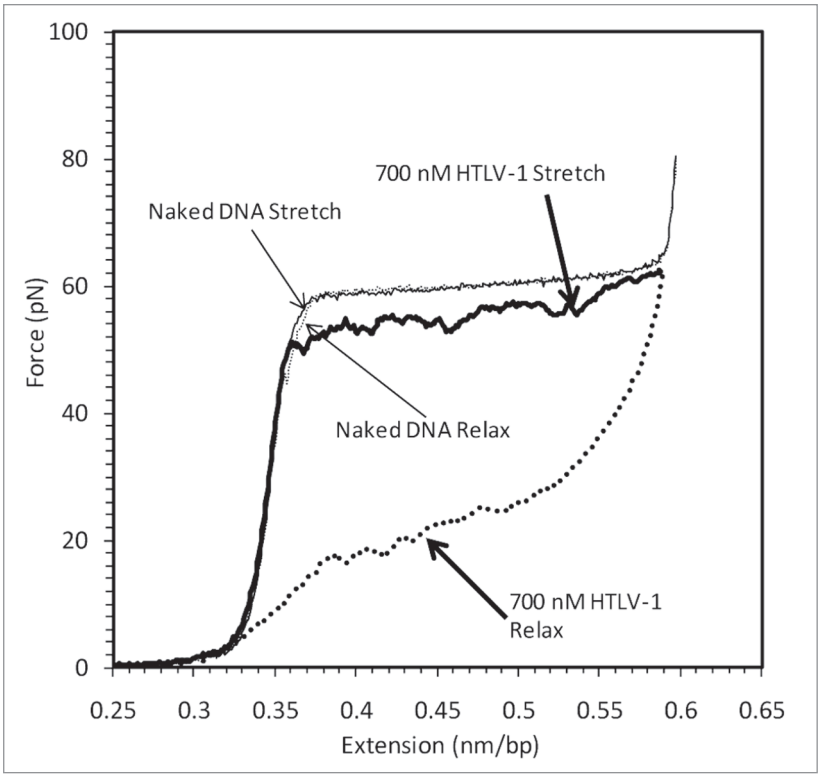

Figure 5. DNA stretching (solid line) and relaxation (dotted line) curves in the absence of protein (thin black line) and in the presence of 700 nM HTLV-1 NC (thick black line).

activity of NC.[78] This can be more informative than other more involved methods for studying rapid protein-NA interaction kinetics, such as stop-flow techniques. Thus, for example, the major difference between the efficient NA chaperone NC protein and the SSB proteins discussed earlier becomes immediately apparent from the comparison of the proteins' effects on the DNA stretching hysteresis (compare data in **Figs. 3 and 4**). Indeed, in contrast to the "fast" NC proteins, the SSB proteins tend to stay "permanently" on ssDNA on the few minute time scale of our experiment.

The optimal NA chaperone activity of HIV-1 NC becomes further apparent by comparison with NC proteins from other retroviruses. NC proteins from Rous sarcoma virus (RSV), Moloney murine leukemia virus (MLV) and human T-cell lymphotropic virus type 1 (HTLV-1) were examined using DNA stretching and other methods.[95] From their force extension curves, it was shown that the presence of RSV NC leads to an increase in the overall DNA melting force, similar to that observed for multivalent cations, indicating that RSV may slightly stabilize NA. Very little hysteresis was observed, showing that RSV has the rapid DNA binding kinetics observed for HIV-1 NC. However, it lacks HIV-1 NC's significant duplex destabilizing ability. In contrast, MLV does not change the overall melting force as much as HIV-1 NC, but it does exhibit some DNA duplex destabilization. The presence of MLV induces moderate hysteresis, similar to the SSB T7 gp2.5. Both RSV and MLV NC exhibit only some of the components of NA chaperone activity, while missing others. Importantly, the rapid kinetics of the NC-DNA interaction correlated best with the protein's ability to facilitate the annealing kinetics of complementary TAR RNA and DNA hairpins measured in the same study.[95] At the same time, these NC proteins' ability to destabilize DNA duplex appeared to be rather similar for all tested NC proteins, including the worst chaperone—HTLV-1 NC.[95-97]

Although some defects in the ability of mutant NC proteins to facilitate NA annealing were consistently observed,[67,92,95,96,98,99] these defects are not nearly as strong as the up to 6 orders of magnitude decrease in viral infectivity measured in vivo for the viruses with the same mutations in NC.[100-102] While simple sedimentation assays suggest that these mutant NC proteins are still capable of aggregating NA,[41,51,95,96] it is possible that these "slow" NC proteins may interfere with reverse transcription by partially losing their ability to tightly aggregate NA within the capsid, thereby leading to early capsid uncoating prior to the completion of the RTion. The latter can occur due to the increase in gRNA volume within the virion, leading to excessive pressure on the capsid and capsid uncoating. Such early capsid uncoating can become detrimental for the virus, as all of the proteins associated with the reverse transcription complex will, most likely, be lost prior to completion of RTion.

In the extreme case of HTLV-1 NC, its complete inability to facilitate TAR RNA-DNA annealing correlates with the fact that it exhibits the largest DNA stretching hysteresis relative to other NC proteins, as shown in **Figure 5**.[95,96] This NC protein is also known to be unable to aggregate NA. The slow off rate of HTLV-1 NC protein from ssDNA can be quantitatively characterized by measuring the relaxation time of the force, as the DNA length is held constant at some point during the DNA release after its force-induced melting. Previous results show that the off rates of T4 gp32 and T7gp2.5 are on the order of a few minutes and tens of seconds respectively.[71,76] Similar experiments on HTLV-1 NC indicate that this protein has a slow off rate on the same order as T4 gp32, with a time constant of about 50 s. The same time decay experiment was also performed on a C-terminal truncation mutant, HTLV-1 NCΔC29, in which 29 C-terminal residues are removed.[96] Interestingly, truncating the C-terminus increases the off rate ~10-fold, with a time constant of about 5 s, similar to the off rate of T7 gp2.5. Thus, a model in which the C-terminal domain of HTLV-1 NC regulates its NA interactions (reminiscent of similar models describing T4 gp32 and T7 gp2.5), is proposed to explain this result. For all three of these proteins (T7 gp2.5, T4 gp32 and HTLV-1 NC), prior to binding DNA, their cationic nucleic acid binding sites are occupied by some portion of the C-terminal domain (CTD), which electrostatically mimics ssDNA. The difference is that the NA binding sites of T4 gp32 and HTLV-1 NC interact with their own CTD, rather than their partners' CTD, which is the mechanism of T7 gp2.5. In either case, truncating the CTD exposes the NA binding sites, facilitates protein-DNA binding and increases

the on rate. In addition, HTLV-1 NC binds to ssDNA with at least moderate cooperativity with the participation of the CTD. Thus, deleting HTLV-1 NC's CTD also increases the off rate. In contrast, T7 gp2.5 binds with little or no cooperativity. For T4 gp32, it does bind with cooperativity, but the cooperativity is salt-independent and T4 gp32's CTD is not involved in the protein-DNA interaction. Therefore, the removal of the CTD of T4 gp32 only increases its DNA on rate, leaving the off rate unchanged. Thus, while these three proteins all exhibit regulated DNA interactions, the precise biophysical mechanism of the interaction differs in each case.

As a retroviral NC protein, HTLV-1 NC does not exhibit the rapid kinetics of other NC proteins. Instead, from a kinetics perspective, it behaves more like an SSB protein. The advantage of this property is still unknown, but a possible proposal is that HTLV-1 has sacrificed nucleic acid chaperone potency for other biological purposes, such as coating ssDNA to prevent packaging of APOBEC3G.[103] Incidentally, the anionic CTD of HTLV-1 NC appears to be rather similar in amino acid composition and charge to the p6 domain of the HIV-1 NC precursor protein HIV-1 NCp15.[96] Moreover, both proteins share the property of slow ssDNA interaction kinetics,[41,96] the inability to facilitate NA aggregation,[21,96,104] and poor NA annealing activity.[96] The biological role of this coincidence is still unknown.

NA Chaperone Proteins: LINE-1 ORF1p

Retrotransposons are genetic elements that can amplify themselves in a genome and are ubiquitous components of the DNA of many eukaryotic organisms. According to the structure and mechanism, retrotransposons can be divided into two types, long terminal repeat (LTR) retrotransposons, with direct LTRs that range from ~100 bp to over 5 kb and non-LTR retrotransposons. The retroviruses mentioned above use LTR regions of genomic RNA for minus-strand transfer, therefore, they are similar to LTR retrotransposons. Some other retrotransposons, like long interspersed nuclear element type 1 (LINE-1 or L1), which also reproduce via an RNA intermediate but lack this repeated region, belong to non-LTR retrotransposons.[105] L1 encodes the products of its two open reading frames, ORF1 and ORF2, which are essential for retrotransposition. The similarity of its sequence to retroviral reverse transcriptases predicted that ORF2p provides both the endonuclease and reverse transcriptase activities. However, the function of ORF1p was unknown until single-molecule experiments were performed to demonstrate its capability to act as a nucleic acid chaperone protein.[106]

Figure 6 shows the force extension curve in the presence of L1 ORF1p. Comparing it to the force extension curve of naked DNA, we find that the presence of L1 ORF1p affects the curve in three aspects. At the beginning of the curve, the force required to stretch DNA to B-form contour length increases. This difference indicates that L1 ORF1p induces large DNA condensation, most likely due to the electrostatic interaction. Secondly, in the presence of ORF1p, the ssDNA to dsDNA transition plateau is dramatically inclined. This increase in the width of the transition force saturates at about 15pN, indicating even low concentrations

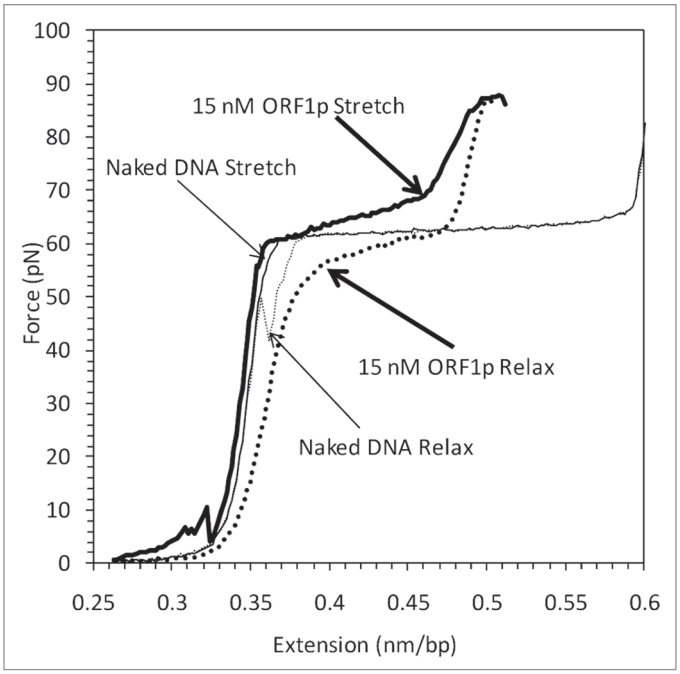

Figure 6. DNA stretching (solid line) and relaxation (dotted line) curves in the absence of protein (thin black line) and in the presence of 15 nM ORF1p (thick black line).

of ORF1p can significantly alter the cooperativity of dsDNA melting. Finally, the relaxation curve is somewhat close to the stretching curve, showing that ORF1p induces moderate hysteresis and does not have rapid kinetics comparable to that of HIV-1 NC, but its kinetics are similar to other NC proteins such as those from RSV and to some extent MLV.[95]

In addition, experiments with L1 ORF1p mutants demonstrate that two arginine (R) residues in the basic C-terminal domain, at positions 297–298, are integral to nucleic acid binding and chaperone activity.[14] Substituting alanine (A) or lysine (K) for one or both of these arginine (R) residues results in a series of mutants, RK, KR, KK and AA. The DNA stretching experiments show significant changes in the presence of these mutants. In contrast, wild type, RK, KR and KK all exhibit similar NA binding affinities, while the AA mutant compromises binding affinity significantly. Surprisingly, however, the mutants with similar binding affinity show very different retrotransposition activity. The property of the mutants that correlates most strongly with retrotransposition is aggregation. The mutants that aggregated DNA too strongly as observed by DNA stretching were not effective at facilitating retrotransposition. This suggests that, although DNA aggregation is a necessary component of NA chaperone activity, too much of this component can be detrimental. This is likely due to the fact that NA can be trapped in specific aggregated states with low NA mobility, thereby unable to reach the final lowest energy states.

Overall, wild type ORF1p exhibits high affinity for ssDNA and strongly aggregates DNA. Therefore, this protein is among the best chaperone proteins known at low concentrations. However, because it is not a good helix-destabilizer, it is not expected to be an optimal chaperone protein for processes

requiring this property, such as minus-strand transfer during HIV-1 reverse transcription. This is consistent with the requirements for LINE-1 retrotransposition, as the poly-A tail already forms a relatively weakly base-paired structure. Specific mutations in ORF1p tend to inhibit its chaperone activity by making it a strong helix-stabilizer and inducing aggregation that is too strong, illustrating the need to balance the components of nucleic acid chaperone activity for optimal function.[52,106]

Conclusions

In this review, we discuss the properties of multivalent cations, SSB proteins, as well as nucleic acid chaperone proteins from retroviruses and retrotransposons revealed using single-molecule biophysical methods. We attempt to identify the characteristics of an optimal NA chaperone. To be an optimal NA chaperone, a protein must be able to promote NA aggregation, destabilize dsDNA and exhibit rapid kinetics simultaneously. These characteristics must be carefully balanced for efficient chaperone activity, and they must also likely be tuned to the specific biological system in which they act. Multivalent cations strongly interact with dsDNA and induce obvious aggregation. However, instead of destabilizing dsDNA, the presence of multivalent cations significantly stabilizes the dsDNA, increasing the melting force. Therefore, they do not represent an optimal chaperone for systems that require destabilization of relatively complex structures that will not fluctuate open on their own. SSB proteins, such as T4 gp32 and T7gp2.5, are able to destabilize the duplex DNA, which is necessary, for example, in the minus-strand transfer process of HIV-1. However, the large hysteresis on the force extension curves indicates slow kinetics, preventing the SSB protein from facilitating annealing under most conditions. Like other NC proteins, HTLV-1 NC has the ability to destabilize duplex DNA. However, in terms of kinetics it behaves most like SSB proteins. The reason why HTLV-1 sacrifices its rapid kinetics is still unknown, but without rapid kinetics, HTLV-1 is not an optimal chaperone. In contrast, the C-terminal truncation mutants of T4gp32, T7gp2.5 and HTLV-1 NC all exhibit faster kinetics, due to the exposure of cationic nucleic acid binding sites. Force extension curves of DNA in the presence of HIV-1 NC demonstrate that HIV-1 NC has high NA affinity, greatly lowers the overall melting force and exhibits rapid dissociation kinetics, satisfying all the requirements for efficient chaperone activity. In addition, ORF1p also has high ssDNA affinity and strongly aggregates DNA, suggesting that it will be a good chaperone for processes in which relatively unstable secondary structures must be rearranged. Specific ORF1p mutants strongly increase aggregation and also inhibit retrotransposition, indicating that a balance of aggregation and ssDNA affinity is needed to optimize NA chaperone activity during LINE-1 retrotransposition.

Acknowledgements

We acknowledge funding from US National Institutes of Health (R01GM072462) and the US National Science Foundation (MCB-0744456).

References

1. Tsuchihashi Z, Brown PO. DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein. J Virol 1994; 68:1-2.
2. Rein A, Henderson LE, Levin JG. Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication. Trends in Biochem Sci 1998; 23:297-301.
3. Marko JF, Siggia ED. Stretching DNA. Macromolecules 1995; 28:8759-70.
4. Smith SB, Cui YJ, Bustamante C. Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules. Science 1996; 271:795-9.
5. Williams MC, Rouzina I. Force Spectroscopy of Single DNA and RNA Molecules. Curr Opin in Struct Biol 2002; 12:330-6.
6. Williams MC, Rouzina I, Bloomfield VA. Thermodynamics of DNA Interactions from Single Molecule Stretching Experiments. Acc Chem Res 2002; 35:159-66.
7. Williams MC, Wenner JR, Rouzina I, Bloomfield VA. Effect of pH on the Overstretching Transition of Double-Stranded DNA: Evidence of Force-Induced DNA Melting. Biophys J 2001; 80:874-81.
8. Williams MC, Wenner JR, Rouzina I, Bloomfield VA. Entropy and Heat Capacity of DNA Melting from Temperature Dependence of Single Molecule Stretching. Biophys J 2001; 80:1932-9.
9. Paramanathan T, Westerlund F, McCauley MJ, Rouzina I, Lincoln P, Williams MC. Mechanically manipulating the DNA threading intercalation rate. J Am Chem Soc 2008; 130:3752-3.
10. Shokri L, Rouzina I, Williams MC. Interaction of bacteriophage T4 and T7 single-stranded DNA-binding proteins with DNA. Phys Biol 2009; 6:25002.
11. Vladescu ID, McCauley MJ, Nunez ME, Rouzina I, Williams MC. Quantifying force-dependent and zero-force DNA intercalation by single-molecule stretching. Nature Methods 2007; 4:517-22.
12. Vladescu ID, McCauley MJ, Rouzina I, Williams MC. Mapping the phase diagram of single DNA molecule force-induced melting in the presence of ethidium. Phys Rev Lett 2005; 95:158102.
13. van Mameren J, Gross P, Farge G, Hooijman P, Modesti M, Falkenberg M, et al. Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging. Proc Natl Acad Sci USA 2009; 106:18231-6.
14. Chaurasiya KR, Paramanathan T, McCauley MJ, Williams MC. Biophysical characterization of DNA binding from single molecule force measurements. Phys Life Rev 2010; 7:358-61.
15. Williams MC, Rouzina I, McCauley MJ. Peeling back the mystery of DNA overstretching. Proc Natl Acad Sci USA 2009; 106:18047-8.
16. Danilowicz C, Limouse C, Hatch K, Conover A, Coljee VW, Kleckner N, et al. The structure of DNA overstretched from the 5'5' ends differs from the structure of DNA overstretched from the 3'3' ends. Proc Natl Acad Sci USA 2009; 106:13196-201.
17. McDonald D, Vodicka MA, Lucero G, Svitkina TM, Borisy GG, Emerman M, et al. Visualization of the intracellular behavior of HIV in living cells. J Cell Biol 2002; 159:441-52.
18. Arhel NJ, Souquere-Besse S, Munier S, Souque P, Guadagnini S, Rutherford S, et al. HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. EMBO J 2007; 26:3025-37.
19. Arfi V, Lienard J, Nguyen XN, Berger G, Rigal D, Darlix JL, et al. Characterization of the behavior of functional viral genomes during the early steps of human immunodeficiency virus type 1 infection. J Virol 2009; 83:7524-35.
20. Warrilow D, Stenzel D, Harrich D. Isolated HIV-1 core is active for reverse transcription. Retrovirology 2007; 4:77.
21. Mirambeau G, Lyonnais S, Coulaud D, Hameau L, Lafosse S, Jeusset J, et al. HIV-1 protease and reverse transcriptase control the architecture of their nucleocapsid partner. PLoS ONE 2007; 2:669.
22. Bloomfield VA. DNA condensation by multivalent cations. Biopolymers 1997; 44:269-82.
23. Bloomfield VA. Condensation of DNA by multivalent cations: considerations on mechanism. Biopolymers 1991; 31:1471-81.
24. Bloomfield VA, Crothers DM, Tinoco I. Nucleic Acids Structure, Properties and Function. Sausalito, CA: University Science Books 2000.
25. Bloomfield VA. DNA condensation by multivalent cations. Biopolymers 1998; 44:269-82.
26. Vilfan ID, Conwell CC, Sarkar T, Hud NV. Time study of DNA condensate morphology: implications regarding the nucleation, growth and equilibrium populations of toroids and rods. Biochemistry 2006; 45:8174-83.
27. Hud NV, Vilfan ID. Toroidal DNA condensates: unraveling the fine structure and the role of nucleation in determining size. Annu Rev Biophys Biomol Struct 2005; 34:295-318.
28. He S, Arscott PG, Bloomfield VA. Condensation of DNA by multivalent cations: experimental studies of condensation kinetics. Biopolymers 2000; 53:329-41.
29. Pelta J, Livolant F, Sikorav JL. DNA condensation by polyamines and cobalthexamine. J Biol Chem 1996; 271:13644-51.
30. Pelta JJ, Durand D, Doucet J, Livolant F. DNA mesophases induced by spermidine: structural properties and biological implications. Biophys J 1996; 71:48-63.
31. Raspaud E, Chaperon I, Leforestier A, Livolant F. Spermine-induced aggregation of DNA, nucleosome and chromatin. Biophys J 1999; 77:1547-55.

32. Kabanov AV, Kabanov VA. DNA complexes with Polycations for the Delivery of Genetic materials into Cells. Bioconjugate Chemistry 1995; 6:7-20.
33. Koculi E, Lee NK, Thirumalai D, Woodson SA. Folding of the Tetrahymena ribozyme by polyamines: importance of counterion valence and size. J Mol Biol 2004; 341:27-36.
34. Koculi E, Thirumalai D, Woodson SA. Counterion charge density determines the position and plasticity of RNA folding transition states. J Mol Biol 2006; 359:446-54.
35. Toma AC, de Frutos M, Livolant F, Raspaud E. DNA condensed by protamine: a "short" or "long" polycation behavior. Biomacromolecules 2009; 10:2129-34.
36. Le Cam E, Coulaud D, Delain E, Petitjean P, Roques BP, Gerard D, et al. Properties and growth mechanism of the ordered aggregation of a model RNA by the HIV-1 nucleocapsid protein: an electron microscopy investigation. Biopolymers 1998; 45:217-29.
37. Stoylov SP, Vuilleumier C, Stoylova E, De Rocquigny H, Roques BP, Gerard D, et al. Ordered aggregation of ribonucleic acids by the human immunodeficiency virus type 1 nucleocapsid protein. Biopolymers 1997; 41:301-12.
38. Krishnamoorthy G, Roques B, Darlix JL, Mély Y. DNA condensation by the nucleocapsid protein of HIV-1: a mechanism ensuring DNA protection. Nucleic Acids Res 2003; 31:5425-32.
39. Vo MN, Barany G, Rouzina I, Musier-Forsyth K. Effect of Mg(2+) and Na(+) on the nucleic acid chaperone activity of HIV-1 nucleocapsid protein: implications for reverse transcription. J Mol Biol 2009; 386:773-88.
40. Vo MN, Barany G, Rouzina I, Musier-Forsyth K. HIV-1 nucleocapsid protein switches the pathway of transactivation response element RNA/DNA annealing from loop-loop "kissing" to "zipper". J Mol Biol 2009; 386:789-801.
41. Cruceanu M, Urbaneja MA, Hixson CV, Johnson DG, Datta SA, Fivash MJ, et al. Nucleic acid binding and chaperone properties of HIV-1 Gag and nucleocapsid proteins. Nucleic Acids Res 2006; 34:593-605.
42. Sikorav JL, Church GM. Complementary recognition in condensed DNA: accelerated DNA renaturation. J Mol Biol 1991; 222:1085-108.
43. Rouzina I, Bloomfield VA. Macroion Attraction Due to Electrostatic Correlation between Screening Counterions. (1) Mobile Surface-Adsorbed Ions and Diffuse Ion Cloud. J Phys Chem 1996; 100:9977-89.
44. Nguyen TT, Rouzina I, Shklovskii BI. Reentrant Condensation of DNA induced by Multivalent Counterions. J Chem Phys 2000; 112:2562-8.
45. Shklovskii BI, Nguyen TT. Kinetics of macroion coagulation induced by multivalent counterions. Phys Rev E Stat Nonlin Soft Matter Phys 2002; 65:031409-15.
46. Stevens MJ. Simple simulations of DNA condensation. Biophys J 2001; 80:130-9.
47. Baumann CG, Bloomfield VA, Smith SB, Bustamante C, Wang MD, Block SM. Stretching of Single Collapsed DNA Molecules. Biophys J 2000; 78:1965-78.
48. Murayama Y, Sakamaki Y, Sano M. Elastic response of single DNA molecules exhibits a reentrant collapsing transition. Phys Rev Lett 2003; 90:018102.
49. Besteman K, Hage S, Dekker NH, Lemay SG. Role of tension and twist in single-molecule DNA condensation. Phys Rev Lett 2007; 98:058103.
50. van den Broek B, Noom MC, van Mameren J, Battle C, Mackintosh FC, Wuite GJ. Visualizing the formation and collapse of DNA toroids. Biophys J 2010; 98:1902-10.
51. Cruceanu M, Gorelick RJ, Musier-Forsyth K, Rouzina I, Williams MC. Rapid kinetics of protein-nucleic acid interaction is a major component of HIV-1 nucleocapsid protein's nucleic acid chaperone function. J Mol Biol 2006; 363:867-77.
52. Martin SL, Bushman D, Wang F, Li PW, Walker A, Cummiskey J, et al. A single amino acid substitution in ORF1 dramatically decreases L1 retrotransposition and provides insight into nucleic acid chaperone activity. Nucleic Acids Res 2008; 36:5845-54.
53. Zhang R, Shklovskii BI. The pulling force of a single DNA molecule condensed by spermidine Zhang R and Shklovskii BI. Physica A: Statistical and Theoretical Physics 2005; 349:563-70.
54. Frank-Kamenetskii MD, Anshelevich VV, Lukashin AV. Polyelectrolyte model of DNA. <Translation> Soviet Physics-Uspekhi 1987; 151:595-618.
55. Baumann CG, Smith SB, Bloomfield VA, Bustamante C. Ionic Effects On the Elasticity Of Single DNA Molecules. Proc Natl Acad Sci USA 1997; 94:6185-90.
56. Rouzina I, Bloomfield VA. DNA Bending by Small, Mobile Multivalent Cations. Biophys J 1998; 74:3152-64.
57. Williams MC, Rouzina I, Wenner JR, Gorelick RJ, Musier-Forsyth K, Bloomfield VA. Mechanism for nucleic acid chaperone activity of HIV-1 nucleocapsid protein revealed by single molecule stretching. Proc Natl Acad Sci USA 2001; 98:6121-6.
58. Azoulay J, Clamme JP, Darlix JL, Roques BP, Mely Y. Destabilization of the HIV-1 Complementary Sequence of TAR by the Nucleocapsid Protein Through Activation of Conformational Fluctuations. J Mol Biol 2003; 326:691-700.
59. Bernacchi S, Stoylov SP, Piemont E, Ficheux D, Roques B, Darlix JL, et al. HIV-1 nucleocapsid protein activates transient melting of least stable parts of the secondary structure of TAR and its complementary sequence. J Mol Biol 2002; 317:385-99.
60. Beltz H, Clauss C, Piemont E, Ficheux D, Gorelick RJ, Roques B, et al. Structural determinants of HIV-1 nucleocapsid protein for cTAR DNA binding and destabilization and correlation with inhibition of self-primed DNA synthesis. J Mol Biol 2005; 348:1113-26.
61. Frank-Kamenetskii F. Simplification of the empirical relationship between melting temperature of DNA, its GC content and concentration of sodium ions in solution. Biopolymers 1971; 10:2623-4.
62. Wenner JR, Williams MC, Rouzina I, Bloomfield VA. Salt Dependence of the Elasticity and Overstretching Transition of Single DNA Molecules. Biophys J 2002; 82:3160-9.
63. Hargittai MRS, Gorelick RJ, Rouzina I, Musier-Forsyth K. Mechanistic insights into the kinetics of HIV-1 Nucleocapsid Protein-facilitated tRNA annealing to the Primer Binding Site. J Mol Biol 2004; 337:951-68.
64. Cosa G, Harbron EJ, Zeng Y, Liu HW, O'Connor DB, Eta-Hosokawa C, et al. Secondary structure and secondary structure dynamics of DNA hairpins complexed with HIV-1 NC protein. Biophys J 2004; 87:2759-67.
65. Lee N, Gorelick RJ, Musier-Forsyth K. Zinc finger-dependent HIV-1 nucleocapsid protein-TAR RNA interactions. Nucleic Acids Res 2003; 31:4847-55.
66. Levin JG, Guo J, Rouzina I, Musier-Forsyth K. Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism. Prog Nucleic Acid Res Mol Biol 2005; 80:217-86.
67. Guo J, Wu T, Anderson J, Kane BF, Johnson DG, Gorelick RJ, et al. Zinc Finger Structures in the Human Immunodeficiency Virus Type 1 Nucleocapsid Protein Facilitate Efficient Minus- and Plus-Strand Transfer. J Virol 2000; 74:8980-8.
68. Hargittai MR, Mangla AT, Gorelick RJ, Musier-Forsyth K. HIV-1 nucleocapsid protein zinc finger structures induce tRNA(Lys3) structural changes but are not critical for primer/template annealing. J Mol Biol 2001; 312:985-97.
69. Piche C, Schernthaner JP. Optimization of in vitro transcription and full-length cDNA synthesis using the T4 bacteriophage gene 32 protein. J Biomol Tech 2005; 16:239-47.
70. Anthony RM, Destefano JJ. In vitro synthesis of long DNA products in reactions with HIV-RT and nucleocapsid protein. J Mol Biol 2007; 365:310-24.
71. Pant K, Karpel RL, Williams MC. Kinetic Regulation of Single DNA Molecule Denaturation by T4 Gene 32 Protein Structural Domains. J Mol Biol 2003; 327:571-8.
72. Pant K, Karpel RL, Rouzina I, Williams MC. Mechanical measurement of single-molecule binding rates: kinetics of DNA helix-destabilization by T4 gene 32 protein. J Mol Biol 2004; 336:851-70.
73. Pant K, Karpel RL, Rouzina I, Williams MC. Salt dependent binding of T4 gene 32 protein to single and double-stranded DNA: single molecule force spectroscopy measurements. J Mol Biol 2005; 349:317-30.
74. Rouzina I, Pant K, Karpel RL, Williams MC. Theory of electrostatically regulated binding of T4 gene 32 protein to single- and double-stranded DNA. Biophys J 2005; 89:1941-56.
75. Pant K, Shokri L, Karpel RL, Morrical SW, Williams MC. Modulation of T4 gene 32 protein DNA binding activity by the recombination mediator protein UvsY. J Mol Biol 2008; 380:799-811.
76. Shokri L, Marintcheva B, Richardson CC, Rouzina I, Williams MC. Single molecule force spectroscopy of salt-dependent bacteriophage T7 gene 2.5 protein binding to single-stranded DNA. J Biol Chem 2006; 281:38689-96.
77. Shokri L, Marintcheva B, Eldib M, Hanke A, Rouzina I, Williams MC. Kinetics and thermodynamics of salt-dependent T7 gene 2.5 protein binding to single- and double-stranded DNA. Nucleic Acids Res 2008; 36:5668-77.
78. Cruceanu M, Stephen AG, Beuning PJ, Gorelick RJ, Fisher RJ, Williams MC. Single DNA molecule stretching measures the activity of chemicals that target the HIV-1 nucleocapsid protein. Anal Biochem 2006; 358:159-70.
79. Driscoll MD, Hughes SH. 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 sequences. J Virol 2000; 74:8785-92.
80. Heilman-Miller SL, Wu T, Levin JG. Alteration of nucleic acid structure and stability modulates the efficiency of minus-strand transfer mediated by the HIV-1 nucleocapsid protein. J Biol Chem 2004; 279:44154-65.
81. Hu WS, Rhodes T, Dang Q, Pathak V. Retroviral recombination: review of genetic analyses. Front Biosci 2003; 8:143-55.
82. Negroni M, Buc H. Mechanisms of retroviral recombination. Ann Rev Genet 2001; 35:275-302.
83. Negroni M, Buc H. Retroviral recombination: what drives the switch? Nat Rev Mol Cell Biol 2001; 2:151-5.
84. Negroni M, Buc H. Recombination during reverse transcription: an evaluation of the role of the nucleocapsid protein. J Mol Biol 1999; 286:15-31.
85. Gao L, Balakrishnan M, Roques BP, Bambara RA. Insights into the multiple roles of pausing in HIV-1 reverse transcriptase-promoted strand transfers. J Biol Chem 2007; 282:6222-31.
86. DeStefano JJ, Bambara RA, Fay PJ. The mechanism of human immunodeficiency virus reverse transcriptase-catalyzed strand transfer from internal regions of heteropolymeric RNA templates. J Biol Chem 1994; 269:161-8.
87. Chen Y, Balakrishnan M, Roques BP, Bambara RA. Acceptor RNA cleavage profile supports an invasion mechanism for HIV-1 minus strand transfer. J Biol Chem 2005; 280:14443-52.
88. Roda RH, Balakrishnan M, Hanson MN, Wohrl BM, Le Grice SF, Roques BP, et al. Role of the Reverse Transcriptase, Nucleocapsid Protein and Template Structure in the Two-step Transfer Mechanism in Retroviral Recombination. J Biol Chem 2003; 278:31536-46.

89. Rein A, Henderson LE, Levin JG. Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication. Trends Biochem Sci 1998; 23:297-301.
90. Rodriguez-Rodriguez L, Tsuchihashi Z, Fuentes GM, Bambara RA, Fay PJ. Influence of human immunodeficiency virus nucleocapsid protein on synthesis and strand transfer by the reverse transcriptase in vitro. J Biol Chem 1995; 270:15005-11.
91. Kanevsky I, Chaminade F, Ficheux D, Moumen A, Gorelick R, Negroni M, et al. Specific interactions between HIV-1 nucleocapsid protein and the TAR element. J Mol Biol 2005; 348:1059-77.
92. Williams MC, Gorelick RJ, Musier-Forsyth K. Specific zinc finger architecture required for HIV-1 nucleocapsid protein's nucleic acid chaperone function. Proc Natl Acad Sci USA 2002; 99:8614-9.
93. Wetmur JG, Davidson N. Kinetics of renaturation of DNA. J Mol Biol 1968; 31:349-70.
94. Wetmur JG. DNA probes: applications of the principles of nucleic acid hybridization. Crit Rev Biochem Mol Biol 1991; 26:227-59.
95. Stewart-Maynard KM, Cruceanu M, Wang F, Vo MN, Gorelick RJ, Williams MC, et al. Retroviral nucleocapsid proteins display nonequivalent levels of nucleic acid chaperone activity. J Virol 2008; 82:10129-42.
96. Qualley DF, Stewart-Maynard KM, Wang F, Mitra M, Gorelick RJ, Rouzina I, et al. 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 2010; 285:295-307.
97. Darugar Q, Kim H, Gorelick RJ, Landes C. Human T-cell lymphotropic virus type 1 nucleocapsid protein-induced structural changes in transactivation response DNA hairpin measured by single-molecule fluorescence resonance energy transfer. J Virol 2008; 82:12164-71.
98. Guo J, Wu T, Kane BF, Johnson DG, Henderson LE, Gorelick RJ, et al. 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 2002; 76:4370-8.
99. Vo MN, Barany G, Rouzina I, Musier-Forsyth K. Mechanistic studies of mini-TAR RNA/DNA annealing in the absence and presence of HIV-1 nucleocapsid protein. J Mol Biol 2006; 363:244-61.
100. Gorelick R, Gagliardi T, Bosche W, Wiltrout T, Coren L, Chabot D, et al. Strict conservation of the retroviral nucleocapsid protein zinc finger is strongly influenced by its role in viral infection processes: Characterization of HIV-1 particles containing mutant nucleocapsid zinc-coordinating sequences. Virology 1999; 256:92-104.
101. Buckman JS, Bosche WJ, Gorelick RJ. Human immunodeficiency virus-type 1 nucleocapsid Zn^{2+}-fingers are required for efficient reverse transcription, initial integration processes and protection of newly synthesized viral DNA. J Virol 2003; 77:1469-80.
102. Thomas JA, Bosche WJ, Shatzer TL, Johnson DG, Gorelick RJ. Mutations in human immunodeficiency virus type 1 nucleocapsid protein zinc fingers cause premature reverse transcription. J Virol 2008; 82:9318-28.
103. Derse D, Hill SA, Princler G, Lloyd P, Heidecker G. Resistance of human T cell leukemia virus type 1 to APOBEC3G restriction is mediated by elements in nucleocapsid. Proc Natl Acad Sci USA 2007; 104:2915-20.
104. Mirambeau G, Lyonnais S, Coulaud D, Hameau L, Lafosse S, Jeusset J, et al. Transmission electron microscopy reveals an optimal HIV-1 nucleocapsid aggregation with single-stranded nucleic acids and the mature HIV-1 nucleocapsid protein. J Mol Biol 2006; 364:496-511.
105. Malik HS, Burke WD, Eickbush TH. The age and evolution of non-LTR retrotransposable elements. Mol Biol Evol 1999; 16:793-805.
106. Martin SL, Cruceanu M, Branciforte D, Wai-Lun Li P, Kwok SC, Hodges RS, et al. LINE-1 retrotransposition requires the nucleic acid chaperone activity of the ORF1 protein. J Mol Biol 2005; 348:549-61.