J|A|C|S ARTICLES

Published on Web 10/02/2009

HIV-1 Nucleocapsid Protein Bends Double-Stranded Nucleic Acids

Hui Wang, Yu-Shan Yeh, and Paul F. Barbara*

Center for Nano and Molecular Science and Technology and Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, Texas 78712

Received August 18, 2009; E-mail: p.barbara@mail.utexas.edu

Abstract: The human immunodeficiency virus type-1 (HIV-1) nucleocapsid (NC) protein is believed to be unique among the nucleic acid (NA) binding proteins encoded by this retrovirus in being highly multifunctional and relatively nonsequence-specific. Underlying many of NC's putative functions, including for example its chaperon-like activity for various steps of HIV-1 reverse transcription, is NC's ability to partially melt short double-stranded regions of structured NAs, which is essentially a consequence of NC's general binding preference for single-stranded bases. Herein we report a different, previously undiscovered, mode of NC/NA interaction, i.e., NC-induced sharp bending of short segments of fully duplexed DNA/DNA and DNA/RNA. We use single-molecule fluorescence resonance energy transfer (SM-FRET) in vitro to probe NC-induced NA bending and associated heterogeneous conformational dynamics for model NC/NA complexes. NC-induced NA bending may have important biological roles in the previously reported NC-mediated condensation of duplex proviral DNA in the HIV-1 life cycle.

Introduction

Human immunodeficiency virus type-1 (HIV-1) nucleocapsid (NC) is a small viral protein (55 amino acids in length) composed of a basic N-terminal domain and two CCHC-type zinc fingers.[1−3] NC behaves as a multifunctional viral protein that is believed to play critical roles in several important steps during the retroviral life-cycle from genomic RNA packaging and assembly to chaperone activity for key nucleic acid rearrangements in reverse transcription, e.g., the minus strand transfer.[3,4] The interactions of HIV-1 NC protein with nucleic acids are extremely complicated, and molecular level understanding of these interactions is still challenging. Although HIV-1 NC binds preferentially to single-stranded regions, such as bulges and loops in RNA or DNA hairpins, it also binds strongly to the double-stranded regions of DNA duplexes in a nonsequence-specific manner.[5−8] It has been reported that saturated coating of the double-stranded HIV-1 plasmid DNA with NC provides complete protection of the DNA from nuclease digestion.[5] NC-induced condensation of the plasmid DNA into highly compact structures is believed to be the key mechanism that ensures such DNA protection.[6] It has also been suggested that the condensation of duplex DNA by NC might facilitate the integration of the proviral DNA into the genome of the host cell.[9] However, molecular-level understanding of this NC-DNA system has been hampered by structural complexity and heterogeneity resulting from the formation of large-scale nucleic acid (NA)−protein aggregates.[6,10]

In this paper we begin the molecular-level investigation of NC-induced condensation of fully duplexed NAs by exploring whether the mechanism for NC-induced condensation involves NC-induced sharp NA bending. Sharp bending of double-stranded DNA (dsDNA) is essential for biomolecular processes such as gene regulation in prokaryotes and eukaryotes and is one of the key features of the genomic DNA packaging in eukaryotic nucleosomes and viral capsids.[11−18] Fully duplexed dsDNA has a bending-persistence length of about 50 nm (140−150 bp) in buffered salt solutions.[19−22] Although bending of DNA duplexes over such short lengths is energetically

(1) Frankel, A. D.; Young, J. A. T. *Annu. Rev. Biochem.* **1998**, *67*, 1–25.
(2) Rein, A.; Henderson, L. E.; Levin, J. G. *Trends Biochem. Sci.* **1998**, *23*, 297–301.
(3) Levin, J. G.; Guo, J. H.; Rouzina, I.; Musier-Forsyth, K. *Prog. Nucleic Acid Res. Mol. Biol.* **2005**, *80*, 217–286.
(4) Darlix, J. L.; Lapadat-Tapolsky, M.; de Rocquigny, H.; Roques, B. P. *J. Mol. Biol.* **1995**, *254*, 523–537.
(5) Lapadat-Tapolsky, M.; de Rocquigny, H.; Van Gent, D.; Roques, B. P.; Plasterk, R.; Darlix, J. L. *Nucleic Acids Res.* **1993**, *21*, 831–839.
(6) Krishnamoorthy, G.; Roques, B.; Darlix, J. L.; Mely, Y. *Nucleic Acids Res.* **2003**, *31*, 5425–5432.
(7) Urbaneja, M. A.; Wu, M.; Casas-Finet, J. R.; Karpel, R. L. *J. Mol. Biol.* **2002**, *318*, 749–764.
(8) Williams, M. C.; Rouzina, I.; Wenner, J. R.; Gorelick, R. J.; Musier-Forsyth, K.; Bloomfield, V. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 6121–6126.
(9) Carteau, S.; Gorelick, R. J.; Bushman, F. D. *J. Virol.* **1999**, *73*, 6670–6679.
(10) Mirambeau, G.; Lyonnais, S.; Coulaud, D.; Hameau, L.; Lafosse, S.; Jeusset, J.; Justome, A.; Delain, E.; Gorelick, R. J.; Le Cam, E. *J. Mol. Biol.* **2006**, *364*, 496–511.
(11) Gralla, J. D. *Cell* **1991**, *66*, 415–418.
(12) Schleif, R. *Annu. Rev. Biochem.* **1992**, *61*, 199–223.
(13) Blackwood, E. M.; Kadonaga, J. T. *Science* **1998**, *281*, 60–63.
(14) Kornberg, R. D.; Lorch, Y. L. *Cell* **1999**, *98*, 285–294.
(15) Jiang, W.; Chang, J.; Jakana, J.; Weigele, P.; King, J.; Chiu, W. *Nature* **2006**, *439*, 612–616.
(16) Lander, G. C.; Tang, L.; Casjens, S. R.; Gilcrease, E. B.; Prevelige, P.; Poliakov, A.; Potter, C. S.; Carragher, B.; Johnson, J. E. *Science* **2006**, *312*, 1791–1795.
(17) Garcia, H. G.; Grayson, P.; Han, L.; Inamdar, M.; Kondev, J.; Nelson, P. C.; Phillips, R.; Widom, J.; Wiggins, P. A. *Biopolymers* **2007**, *85*, 115–130.
(18) Richmond, T. J.; Davey, C. A. *Nature* **2003**, *423*, 145–150.
(19) Hagerman, P. J. *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 265–286.

10.1021/ja9070046 CCC: $40.75

unfavorable, sharp DNA bending over nanometer length scales occurs spontaneously in vivo and in vitro with the help of DNA-bending proteins upon the formation of DNA−protein complexes.[11−18]

Single-molecule spectroscopy has been proven to be a powerful approach for characterizing protein and DNA/RNA conformational dynamics and shedding light on the underlying heterogeneity in complicated biological processes.[23−29] Here we employ single-molecule fluorescence resonance energy transfer (SM-FRET) to examine whether NC can induce bending of a short DNA segment (59-bp) of the proviral DNA corresponding to the transcripts of the transactivation-responsive region (TAR) of the HIV-1 genome. The two ends of the duplex are labeled with Cy3 and Cy5 as the donor and acceptor dyes, respectively. Our results clearly show that the binding of NC to this DNA segment leads to the formation of sharply bent DNA with end-to-end distances within 10 nm (the Förster radius of FRET between Cy3 and Cy5 is ∼6 nm), a length regime that is accessible with FRET detection.[30] The SM-FRET data also show that the bending/unbending dynamics of NC-DNA complexes are highly complex and heterogeneous. Since the protocol for these experiments includes isolated and immobilized single DNA duplexes, complications due to the formation of large-scale nucleic acid−protein aggregates[6,10] are avoided.

Experimental Section

Sample Preparation. DNA and RNA oligonucleotides containing appropriate dye-labeling and biotin-functionalization were purchased from Trilink BioTechnologies (San Diego, CA) and were purified by the supplier using HPLC. All the molecular constructs used in the present studies are listed in Table S1 in Supporting Information. The oligonucleotides were labeled with either Cy3 (donor dye), Cy5 (acceptor dye), or both dyes. Biotin functionalization was added to the nucleotides for the immobilization of the molecules on coverslips through biotin−streptavidin interactions. A TTT or T overhang was added to the DNA sequences to prevent the undesirable G residue quenching effects. HIV-1 NC protein for these experiments was prepared by solid-phase synthesis as described previously.[31,32]

Flow Cell System for SM-FRET Experiments. The reactions were carried out in a home-built flow cell.[32−34] Typically, the nonlabeled or Cy5-labeled oligonucleotides, buffer solutions, and NC protein solutions were selectively flowed into the flow cell to react with the dual-dye-labeled or Cy3-labeled oligonucleotides immobilized on the coverslip surface. The annealing experiments reported herein were carried out with cTAR DNA or TAR RNA concentration at 5 nM, a concentration that is low enough to ensure the suppression of the NC/NA coaggregate formation.[34] It has previously been shown, furthermore, that NC does not aggregate with itself in solution over the entire range of concentrations used herein.[10] The commercial coverslips (Fisher Scientific) were cleaned with piranha (sulfuric acid:hydrogen peroxide, 7:3) and then treated with Vectabond/acetone 1% w/v solutions (Vector Laboratories, Burlingame, CA) for 5 min. Each coverslip was subsequently PEGylated and biotinylated, after which a reaction chamber with inlet and outlet ports (Nanoport, Upchurch Scientific, Oak Harbor, WA) was assembled. The details of the chamber assembly process have been described in earlier work.[32] Each chamber was treated with streptavidin (Molecular Probes, Eugene, OR; 0.2 mg/mL in 25 mM HEPES buffer), followed by the immobilization of the oligonucleotides. The DNA or RNA oligonucleotides were renatured by incubating 500 nM oligonucleotides in HEPES buffer (25 mM HEPES, pH 7.3, 40 mM NaCl) for 3 min at 85 °C, 5 min at 60 °C, and 10 min at 0 °C. Then the solutions of biotin-functionalized oligonucleotide were further diluted to a final concentration of 500 pM in HEPES buffer containing 10 mM $MgCl_2$. The immobilization of oligonucleotides onto the coverslips was accomplished by incubating renatured oligonucleotides in the chamber over the streptavidin-functionalized coverslip surface several times. The individual hairpins immobilized on the coverslips were separated on average by ∼1.5 μm, which is far enough to ensure that there is no intermolecular interaction between the hairpins. All the annealing and DNA bending reactions were carried out in the reaction chamber at room temperature in reaction buffer in an oxygen scavenger system[35] containing β-D-(+)glucose 3% w/v (Sigma-Aldrich, St. Louis, MO), glucose oxidase 0.1 mg/mL, catalase 0.02 mg/mL (Roche Applied Science, Hague Road, IN), and 1% v/v 2-mercaptoethanol (Sigma-Aldrich, St. Louis, MO). The reaction buffer was composed of 25 mM HEPES (pH = 7.3), 40 mM NaCl, and 1 mM $MgCl_2$.

Data Collection and Analysis. A home-built sample scanning confocal optical/data collection system based on a Zeiss inverted microscope[34] was used in these SM-FRET experiments. The sample flow-cell was scanned by a Queensgate *X*,*Y* scanning stage (NPW-XY-100A, Queensgate, Torquay, U.K.). A high numerical aperture, oil immersion microscope objective (Zeiss Fluar, 100×, NA 1.3) was used for excitation (514 nm or 633 nm) and signal collection. The donor and acceptor fluorescence were separated by a dichroic beam splitter (Chroma 650 DCXR, Chroma Tech., VT) into two beams, and each was detected by an avalanche photodiode (APD) (Perkin-Elmer Optoelectronics SPCM-AQR-15, Vaudreuil, QC, Canada).

In our SM-FRET assays, we recorded the time-resolved FRET trajectories of individual dye-labeled molecules immobilized on the coverslip surface using a confocal microscope in two modes: the *image scanning mode* and the *individual trajectory mode*. We use a combination of the image scanning mode and the individual trajectory mode to resolve the conformational dynamics of the NA duplexes over multiple time scales from milliseconds to minutes.

In the image scanning mode, SM-FRET data were collected synchronously through separate detection channels for Cy3 and Cy5 fluorescence intensities while rapidly switching the laser excitation between 514 and 633 nm, which selectively excited Cy3 and Cy5, respectively. For each scanned sample region, three confocal images were obtained simultaneously including a donor channel image (514 nm excitation), an acceptor channel image (514 nm excitation), and a red channel image (633 nm excitation). The SM-FRET images were acquired at several times, *t*, during the course of annealing

(20) Crothers, D. M.; Drak, J.; Kahn, J. D.; Levene, S. D. *Methods Enzymol.* **1992**, *212*, 3–29.
(21) Shimada, J.; Yamakawa, H. *Macromolecules* **1984**, *17*, 689–698.
(22) Zhang, Y. L.; Crothers, D. M. *Biophys. J.* **2003**, *84*, 136–153.
(23) Zhuang, X. W.; Kim, H.; Pereira, M. J. B.; Babcock, H. P.; Walter, N. G.; Chu, S. *Science* **2002**, *296*, 1473–1476.
(24) Myong, S.; Rasnik, I.; Joo, C.; Lohman, T. M.; Ha, T. *Nature* **2005**, *437*, 1321–1325.
(25) Weiss, S. *Nat. Struct. Biol.* **2000**, *7*, 724–729.
(26) Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature* **2002**, *419*, 743–747.
(27) Lu, Q.; Lu, H. P.; Wang, J. *Phys. Rev. Lett.* **2007**, *98*, 128105.
(28) Lu, H. P.; Iakoucheva, L. M.; Ackerman, E. J. *J. Am. Chem. Soc.* **2001**, *123*, 9184–9185.
(29) Liu, R. C.; Hu, D. H.; Tan, X.; Lu, H. P. *J. Am. Chem. Soc.* **2006**, *128*, 10034–10042.
(30) Roy, R.; Hohng, S.; Ha, T. *Nature Methods* **2008**, *5*, 507–516.
(31) Liu, H. W.; Cosa, G.; Landes, C. F.; Zeng, Y. N.; Kovaleski, B. J.; Mullen, D. G.; Barany, G.; Musier-Forsyth, K.; Barbara, P. F. *Biophys. J.* **2005**, *89*, 3470–3479.
(32) Cosa, G.; Harbron, E. J.; Zeng, Y. N.; Liu, H. W.; O'Connor, D. B.; Eta-Hosokawa, C.; Musier-Forsyth, K.; Barbara, P. F. *Biophys. J.* **2004**, *87*, 2759–2767.
(33) Landes, C. F.; Zeng, Y. N.; Liu, H. W.; Musier-Forsyth, K.; Barbara, P. F. *J. Am. Chem. Soc.* **2007**, *129*, 10181–10188.
(34) Liu, H. W.; Zeng, Y. N.; Landes, C. F.; Kim, Y. J.; Zhu, Y. J.; Ma, X. J.; Vo, M.-N.; Musier-Forsyth, K.; Barbara, P. F. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 5261–5267.
(35) Ha, T. *Curr. Opin. Struct. Biol.* **2001**, *11*, 287–292.

reactions, and in-house software was built to find molecules, calibrate stage-drift and crosstalk, subtract image background, and calculate the FRET efficiency for each molecule. The advantage of this mode is that each molecule is exposed to excitation lasers for very short period time (~9 ms) during each scan, and thus photobleaching of Cy5 can be effectively eliminated. The corrected donor and acceptor intensities, $I_D(t)$ and $I_A(t)$, respectively, were used to calculate the time trajectory of the apparent FRET efficiency, $E_A(t)$, as in

$$E_A(t) = \frac{I_A(t)}{I_D(t) + I_A(t)}$$

$E_A(t)$ is related to the actual FRET efficiency, $E_{FRET}(t)$, by the inclusion of the dye quantum efficiencies, φ_i, and detector quantum efficiencies, η_i, as in

$$E_{FRET}(t) = \frac{I_A(t)}{I_A(t) + I_D(t)\frac{\varphi_A\eta_A}{\varphi_D\eta_D}}$$

Here we define $\varphi_A\eta_A/\eta_D\varphi_D$ as the γ factor, which corrects the E_A into E_{FRET}. In the case of the current experimental setup, it was determined that η_A/η_D was constant and φ_A/φ_D varies in different environments. It has been experimentally determined that the γ factor for oligonucleotide hairpins in the buffer solution is 1, and the γ factors of hairpins in 800 nM NC solutions, the DNA duplexes in 800 nM NC solutions, and the DNA duplexes in buffer solutions were determined by the relative φ_A/φ_D value under each corresponding experimental condition. The relative *quantum yields of donor and acceptor* in each condition were measured as the relative emission intensities of the dyes normalized to the values of the hairpins in buffer. The measurements of relative donor emission intensities were performed on Cy3-labeled TAR DNA and duplexes of Cy3-labeled TAR-nonlabeled cTAR DNA under 514 nm laser excitation. The measurements of relative acceptor emission intensities were performed on dual dye-labeled TAR DNA and duplexes of dual dye-labeled TAR-nonlabeled cTAR DNA under 633 nm laser excitation. The relative emission intensities of the donor and acceptor under different conditions are summarized in Table S2 in Supporting Information. The actual FRET efficiency was derived from the apparent FRET efficiency using the γ factor correction.

In the individual trajectory mode, the microscope was positioned to focus on an individual immobilized molecule, and trajectories of donor and acceptor emission intensities were recorded with 1 ms time resolution for several seconds until the dyes were photobleached. In this mode, the 633 nm laser was turned off and the FRET trajectories were obtained under 514 nm excitation. In this way, we can use the image scanning mode to globally image the samples and study relatively slow dynamics on the time scale of minutes, and use the individual trajectory mode to focus on one particular molecule each time and study relatively fast dynamics on the time scales from milliseconds to seconds. The raw data of the single-molecule trajectories were recorded with 1 ms time resolution, and the trajectories shown in Figure 4 and Figure 6 were boxcar time averaged, or "smoothed" in a 10 ms bin time. The FRET autocorrelations were calculated with 1 ms time resolution using the following equation:

$$\mathrm{FRET}_{\mathrm{autocorrelation}}(\tau) = \frac{\langle \delta E_{FRET}(t) \times \delta E_{FRET}(t+\tau) \rangle}{\langle E_{FRET} \rangle \times \langle E_{FRET} \rangle}$$

The ensemble SM-FRET histograms were obtained by combining all the spectroscopic occurrences within SM-FRET trajectories from many individual molecules under the same condition. In the ensemble FRET histograms shown in Figure 6, the data from molecules that were photobleached within 2 s were discarded. The ensemble FRET autocorrelation curves were obtained by averaging single-molecule FRET autocorrelations of many individual molecules under the same condition.

Results and Discussion

SM-FRET Evidence for NC-Induced dsDNA Bending. The dsDNA duplex under current investigations is formed though NC-chaperoned annealing[34,36,37] of a nonlabeled cTAR DNA hairpin to a dual dye-labeled TAR DNA hairpin immobilized on a coverslip. As schematically illustrated in Figure 1A, the resulting duplexes show an extended conformation in buffer and are sharply bent into more compact conformations when multiple NC proteins bind cooperatively to the DNA. Figure 1B shows the SM-FRET trajectories (image scanning mode) of 158 molecules found in a 30 μm × 30 μm region and the corresponding mean FRET values as a function of time during different reaction stages. During time period I when the immobilized TAR DNA molecules are exposed to buffer, the DNA molecules fold into hairpin structures, giving FRET values close to 1 because its 5′ and 3′ ends are in close proximity relative to the Förster radius for this FRET donor/acceptor pair, which is ~6 nm. During time period II when 5 nM nonlabeled cTAR and 800 nM NC are coflowed into the reaction chamber, a previously reported NC-chaperoned strand annealing reaction occurs, producing the desired, immobilized TAR−cTAR DNA duplexes.

During time period III when the TAR−cTAR duplex molecules are exposed to 800 nM NC, the FRET of each individual duplex switches rapidly in a range of FRET values form 0.1 to 0.8, indicating the dynamic changes of the end-to-end distance of duplexes in the range of ~8 nm to ~4 nm, respectively. During time period IV when the duplex molecules are exposed to buffer, FRET values of each individual duplex become very close to 0, indicating end-to-end distances larger than the measurable range of FRET. We attribute the significant reduction of the end-to-end distances of the NC-bound DNA to NC-induced bending of the TAR−cTAR DNA duplex. Analogous results were obtained for TAR−cTAR duplexes formed through thermal annealing of cTAR DNA to TAR DNA (in the absence of NC) prior to surface immobilization (data not shown). The ability of NC to induce NA duplex bending is apparently general. For example, we observe that NC can also induce bending of RNA−DNA hybrid duplexes, as shown in Figure 2 for the fully duplexed TAR DNA−TAR RNA complex.

The observed NC-induced decrease in end-to-end distance for the DNA duplex bound with NC is not a simple consequence of fraying the terminal region or overall DNA melting. In previous papers, we have shown that NC does not fray (i.e., melt) the 3′/5′ end of TAR−cTAR DNA duplex by control experiments performed on the duplexes with donor and acceptor dyes in close proximity.[34,38] These experiments have been reproduced in this study, and the results are presented in Supporting Information (see Figure S1). We also have previously shown that the 3′/5′ ends of dual dye-labeled bulge-free DNA hairpins are not frayed upon NC binding (see Figure S2 in Supporting Information). These results indicate that NC does

(36) Guo, J. H.; Wu, T. Y.; Anderson, J.; Kane, B. F.; Johnson, D. G.; Gorelick, R. J.; Henderson, L. E.; Levin, J. G. *J. Virol.* **2000**, *74*, 8980–8988.
(37) Vo, M. N.; Barany, G.; Rouzina, I.; Musier-Forsyth, K. *J. Mol. Biol.* **2006**, *363*, 244–261.
(38) Zeng, Y. N.; Liu, H. W.; Landes, C. F.; Kim, Y. J.; Ma, X. J.; Zhu, Y. J.; Musier-Forsyth, K.; Barbara, P. F. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 12651–12656.

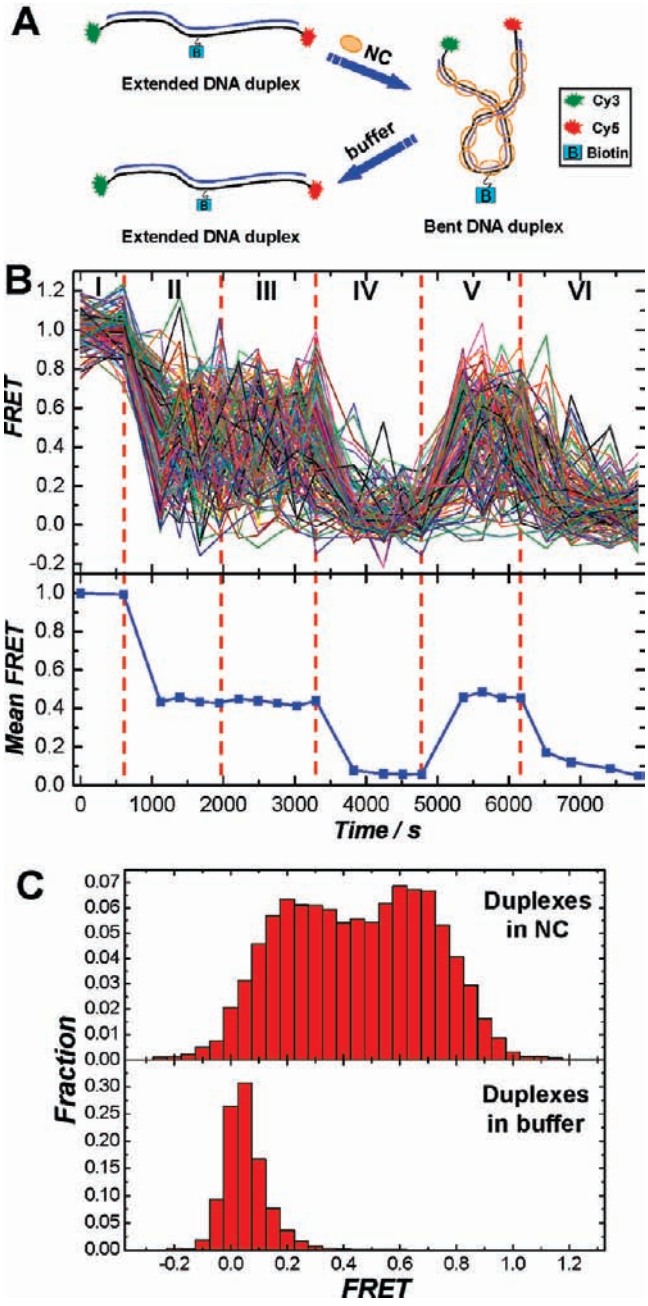

Figure 1. (A) Schematics illustrating the NC-induced sharp bending of TAR−cTAR DNA duplexes. (B) SM-FRET trajectories of 158 molecules found in a 30 μm × 30 μm region (each colored line corresponds to a single molecule) and the corresponding mean FRET values as a function of reaction time. The dual-labeled TAR DNA hairpins are immobilized on a coverslip, and different reagents are flowed into the reaction chamber during different time periods: (I) buffer, (II) 5 nM nonlabeled cTAR DNA and 800 nM NC, (III) 800 nM NC, (IV) buffer, (V) 800 nM NC, and (VI) buffer. (C) FRET histograms of the TAR−cTAR duplex molecules in 800 nM NC (upper) and in buffer (lower).

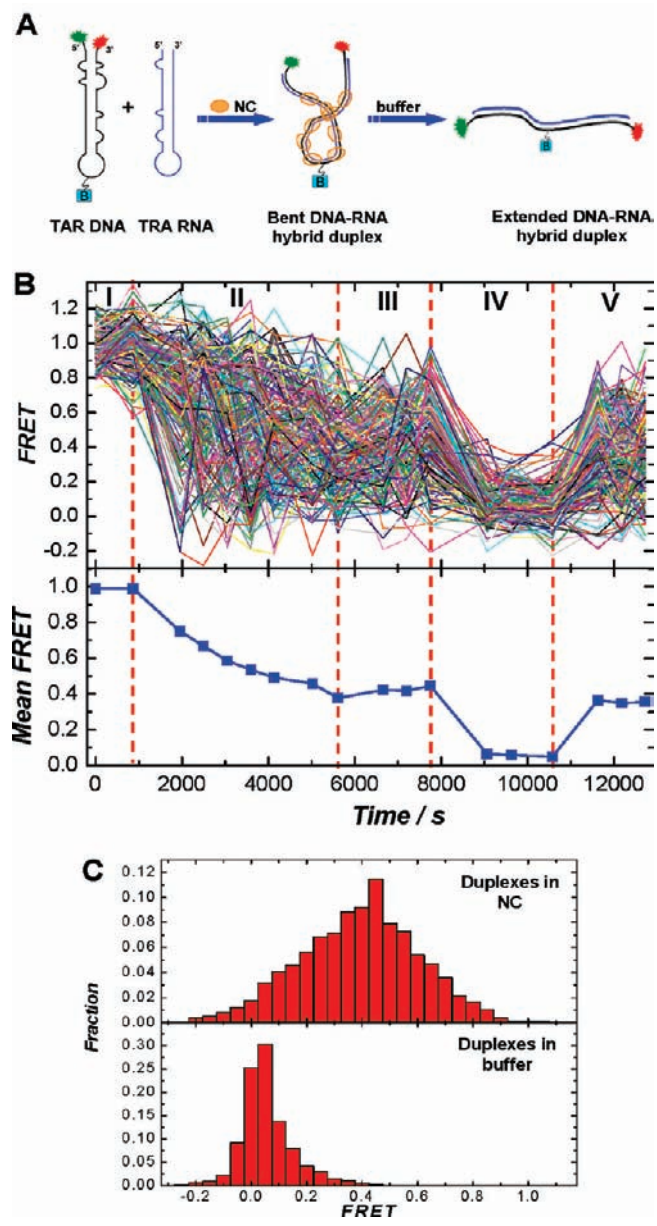

Figure 2. (A) Schematics illustrating the NC-chaperoned annealing of nonlabeled TAR RNA to the immobilized dual-labeled TAR DNA through which the DNA−RNA hybrid duplex is produced. (B) SM-FRET trajectories obtained in the image scanning mode and the corresponding mean FRET values as a function of reaction time. The dual-labeled TAR DNA hairpins are immobilized on a coverslip, and different reagents are flowed into the reaction chamber during different time periods: (I) buffer, (II) 5 nM nonlabeled TAR RNA and 800 nM NC, (III) 800 nM NC, (IV) buffer, and (V) 800 nM NC. (C) FRET histograms of the TAR DNA−TAR RNA hybrid duplex molecules in 800 nM NC (upper) and in buffer (lower).

not significantly melt unstructured (i.e., fully duplexed) DNA sequences of >25 base pairs. This is further supported by the fact that we do not observe reverse annealing of duplex DNA of these lengths in the presence of 800 nM NC, even for periods as long as over 2 h.

Since the dsDNA segments we investigate herein are shorter than DNA's persistence length under these buffer conditions and temperature,[19−22] the naked duplexes should be regarded as relatively rigid structures with end-to-end distances beyond 20 nm, corresponding to FRET close to 0. Our results demonstrate that when bound with NC, the DNA duplexes are spontaneously bent to give smaller end-to-end distances that fall into the detectable regime of our FRET assay. The rapid change of FRET values of individual molecules in Figure 1 clearly indicates dynamical conformational change of the corresponding DNA−NC nucleoprotein complexes. As evident during time period V, the bending of the DNA duplexes is recovered immediately after 800 nM NC is reflowed in, and the duplexes restore their extended conformation during a second buffer wash step (time period VI). This indicates that the NC-induced DNA bending is highly reversible as a result of NC's

reversible binding to DNA. As shown in previous papers, washing NC-hairpin complexes for >100 s returns the FRET histogram data to that for NC free samples.[32] This strongly suggests that NC is removed completely by washing with buffer only. As shown in Figure 1C, the histograms of FRET distribution constructed from thousands of dsDNA−NC complex molecules show a bimodal distribution with two broad peaks centered at ∼0.25 and ∼0.65, respectively, while the naked DNA duplex molecules in buffer only show a single-peaked FRET distribution centered at ∼0.03.

In order to evaluate potential complications in our SM-FRET bending assay due to duplex/surface interactions resulting from the immobilization strategy, we have examined the annealing of 5′-Cy5-labeled cTAR DNA to Cy3-labeled inverted TAR (5′-Cy3 labeled and 3′-biotin functionalized). NC-induced bending is also clearly observed for the duplex formed in this way, as shown in Figure 3, suggesting that surface interaction only plays a minor role in these experiments.

Time-Resolved SM-FRET Measurements. We use a combination of image scanning mode and individual trajectory mode to time-resolve the conformational dynamics of the DNA duplexes over multiple time scales. We use the image scanning mode to globally image the samples and study dynamics on the time scale of minutes or longer, and use the individual trajectory mode to focus on one particular molecule at a time in order to study dynamics on the time scales from milliseconds to several seconds. As shown in Figures 4A−C, the naked DNA duplexes in buffer show very small FRET fluctuations over a broad distribution of time scales from milliseconds to minutes. Figure 4C shows the ensemble SM-FRET histograms obtained from FRET measurements using the single trajectory mode, where ensemble refers to the combination of SM-FRET trajectories from many individual molecules under the same condition. The single-peaked FRET distribution centered at ∼0 corresponds very well to the FRET histograms constructed using the image scanning mode, further verifying the extended conformation of the naked DNA duplexes. This indicates that the equilibrium distribution of naked DNA duplexes lacks a detectable fraction (<3%) of conformations with end-to-end distances comparable or smaller than the Förster radius of Cy3-Cy5 pair (∼6 nm), and furthermore there is no significant evidence of any amplitude of autocorrelation that indicates detectable fluctuation of end-to-end distances on the millisecond or slower time scales. These observations are in agreement with the classic picture of duplex DNA in which dsDNA segments shorter than the persistence length have an extended rod-like conformation. Although some recent experimental observations show that dsDNA might be more locally flexible than described in the classic picture,[39−41] our SM-FRET results apparently show that the sharp bending of TAR−cTAR duplexes in the absence of NC is extremely rare under the conditions of our study.

In striking contrast to the naked dsDNA in buffer solutions, the DNA−NC nucleoprotein complexes show highly bent conformations which interconvert over a broad range of time scales. This is clearly shown in Figure 5, which portrays FRET

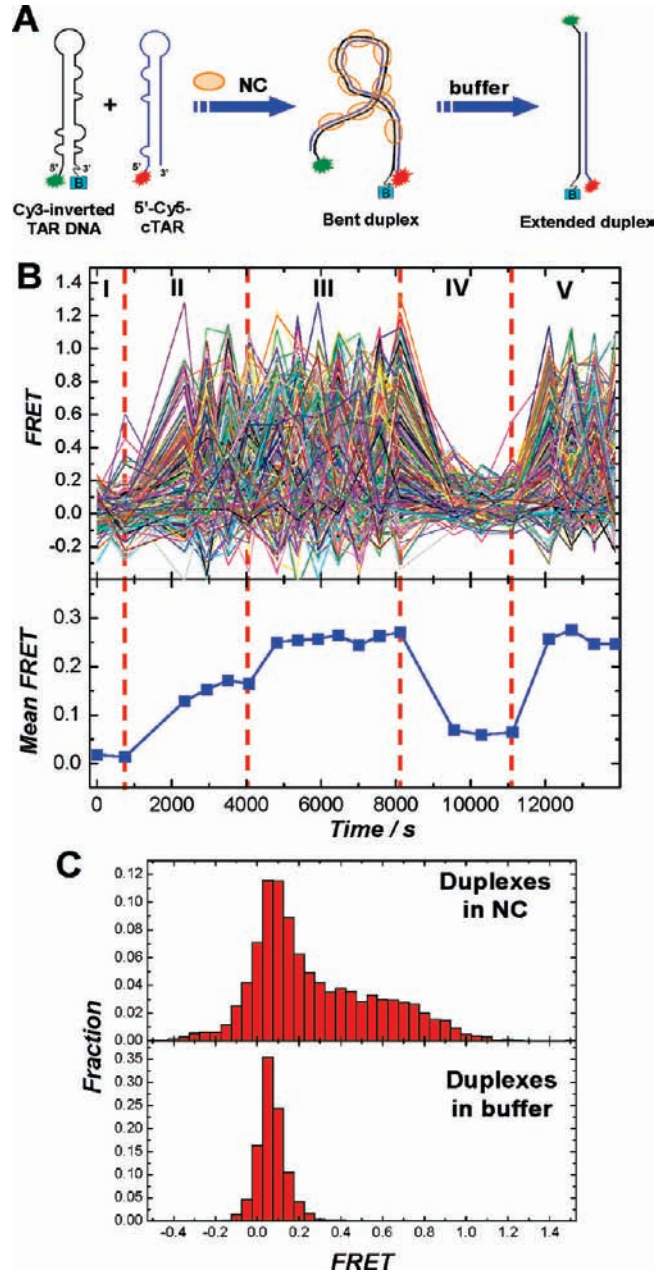

Figure 3. (A) Schematics illustrating the NC-chaperoned annealing of 5′-Cy5-labeled cTAR DNA to the immobilized Cy3-labeled inverted TAR DNA. (B) SM-FRET trajectories obtained in the image scanning mode and the corresponding mean FRET values as a function of reaction time. The Cy3-labeled inverted TAR DNA hairpins are immobilized on a coverslip, and different reagents are flowed into the reaction chamber during different time periods: (I) buffer, (II) 5 nM 5′-Cy5-labeled cTAR DNA and 800 nM NC, (III) 800 nM NC, (IV) buffer, and (V) 800 nM NC. (C) FRET histograms of the DNA duplex molecules in 800 nM NC (upper) and in buffer (lower).

trajectories for an ensemble of 431 dsDNA−NC complexes obtained in the image scanning mode and associated analysis. Figure 5A shows FRET trajectories for the 431 molecules in this ensemble. The corresponding molecularly averaged FRET trajectory is shown in Figure 5B. It is apparent that the trajectories of individual molecules reveal instantaneous FRET values that rapidly switch between high and low values in a fully highly reversible manner. The middle and right columns show a "hole-burning" analysis of the FRET trajectories in which the ensemble of 431 trajectories from the left column

(39) Cloutier, T. E.; Widom, J. *Proc. Natl. Acd. Sci. U.S.A.* **2005**, *102*, 3645–3650.
(40) Mathew-Fenn, R. S.; Das, R.; Harbury, P. A. B. *Science* **2008**, *322*, 446–449.
(41) Wiggins, P. A.; Van der Heijden, T.; Moreno-Herrero, F.; Spakowitz, A.; Phillips, R.; Widom, J.; Dekker, C.; Nelson, P. C. *Nat. Nanotechnol.* **2006**, *1*, 137–141.

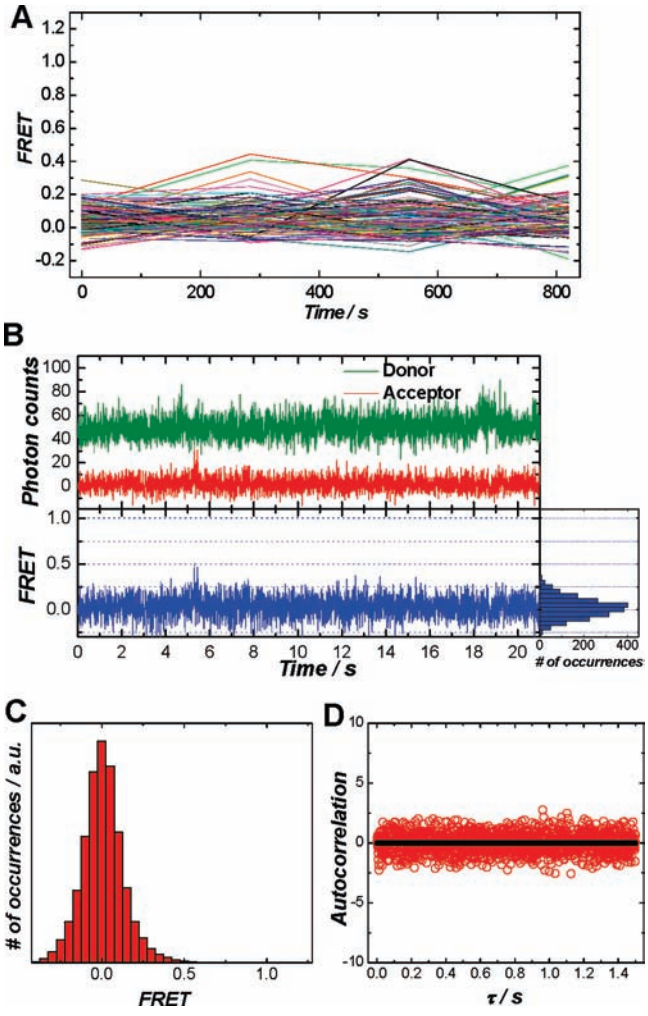

Figure 4. Conformational dynamics of TAR−cTAR DNA duplexes in buffer. (A) SM-FRET trajectories (obtained in the image scanning mode) of 186 molecules found in a 30 μm × 30 μm region. (B) Donor, acceptor, and the corresponding FRET trajectories (obtained in the individual trajectory mode) collected on one representative molecule with a bin time of 10 ms. (C) Ensemble FRET histograms obtained from FRET trajectories of 25 molecules with a bin time of 10 ms. (D) Ensemble FRET autocorrelation obtained from FRET trajectories of 25 molecules with a time resolution of 1 ms.

were sorted into two subensembles of molecules using the FRET value in each trajectory at 1921 s (highlighted by dash lines) and a sorting-FRET threshold of 0.45. The middle column corresponds to the subensemble (205 molecules) with FRET smaller than 0.45 at 1921 s, while the right columns corresponds to the subensemble of 226 moleucles with FRET greater than 0.45, respectively. The trajectories for these subensembles are shown in Figures 5C and 5E, while Figures 5D and 5F portray the molecularly averaged trajectories for these subensembles. The molecularly averaged trajectories clearly show that the interconversion between the high FRET and low FRET conformations are reversible on the time scale of minutes, with a relaxation time that is shorter than the time between two adjacent image frames, i.e., roughly 4 min in this case. Similar results are obtained by performing such "hole-burning" analysis on the images obtained at any of the time spots. The reversible interconversion between the high FRET and low FRET conformations can also be verified by comparing the histograms of all FRET occurrences and of time-averaged FRET. The FRET histograms constructed based on all the occurrences of all molecules in this ensemble (Figure 5G) show the characteristic bimodal distribution, whereas the ensemble FRET histograms of time-averaged FRET of each individual molecule (the averaged FRET of each molecule in all the scanned images) (Figure 5H) show a single-peaked distribution. This further verifies that the interconversion between the high FRET and low FRET conformations is fully reversible on the time scale of minutes.

Although indistinguishable on the time scale of minutes, conformational dynamics on the millisecond to second time scale are obvious in the FRET measurements performed in individual trajectory mode. Figure 6A and 6B show donor and acceptor intensities, FRET trajectories and FRET autocorrelations for representative molecules in the ensemble. The molecule in Figure 6A clearly shows much larger conformational fluctuation than the molecule in Figure 6B on the ~1.5 s time scale of the trajectory measurements. The FRET histograms (Figure 6C) built from a sum of 45 individual molecular trajectories with a dwell time of 10 ms clearly shows a bimodal FRET distribution. The trajectories were roughly sorted into two subensemble using the time averaged FRET for each molecule and a sorting dividing-line of 0.45. The time averaged FRET, denoted as $\langle$FRET$\rangle$, of each molecule refers to the averaged value over all the FRET data points (1 ms time resolution) collected on one particular molecule before the molecule gets photobleached. The subpopulation 1 ($0 < \langle\text{FRET}\rangle < 0.45$) shows much larger end-to-end distance fluctuations and is conformationally more dynamic than the subpopulation 2 ($0.45 < \langle\text{FRET}\rangle < 1$). The subensemble FRET autocorrelation curves of the two subpopulations can be well fit by three-step and two-step exponential decay curves, respectively, with decay times of tens of milliseconds and subseconds. The complicated heterogeneous conformational dynamics suggest the existence of multiple isomers and perhaps a distribution of the number of bound NC for the bent DNA−NC nucleoprotein complexes, as well as the potential existence of multiple interconversion pathways due to the usually observed heterogeneity of NC−nucleic acid interactions. We have also constructed the ensemble FRET histograms and autocorrelation using the FRET trajectories during the very first second of each molecule, and similar bimodal FRET distribution and typical autocorrelation line-shape are well reproduced. Since in our SM-FRET measurements the probability of occurrence of Cy5 photobleaching within 1 s is small (<10%), the complication due to photobleaching or photoblinking of Cy5 can be eliminated by the analysis performed on trajectories during this shorter time period. This verifies that the observed FRET changes are largely due to molecular conformational dynamics without significant interference by photobleaching or photoblinking of Cy5.

NC Concentration Dependence. Studies of NC's effects on the conformational changes of the DNA duplexes as a function of NC concentration (Figure 7) provide further evidence that the bending of DNA is a direct consequence of multiple copies of NC bound to the DNA duplex. As the NC concentration increases, the fraction of bent DNA duplexes vs extended duplexes increases with very little change in the FRET distribution at NC concentrations above 600 nM, due perhaps to a saturated binding of NC to the duplex. The sharp increase in mean FRET values over a narrow NC concentration range strongly indicates that multiple copies of NC are involved in the DNA bending process. We also observe that NC-induced bending of dsDNA can be effectively inhibited when the binding of NC to DNA is blocked by excessive nonspecific binding

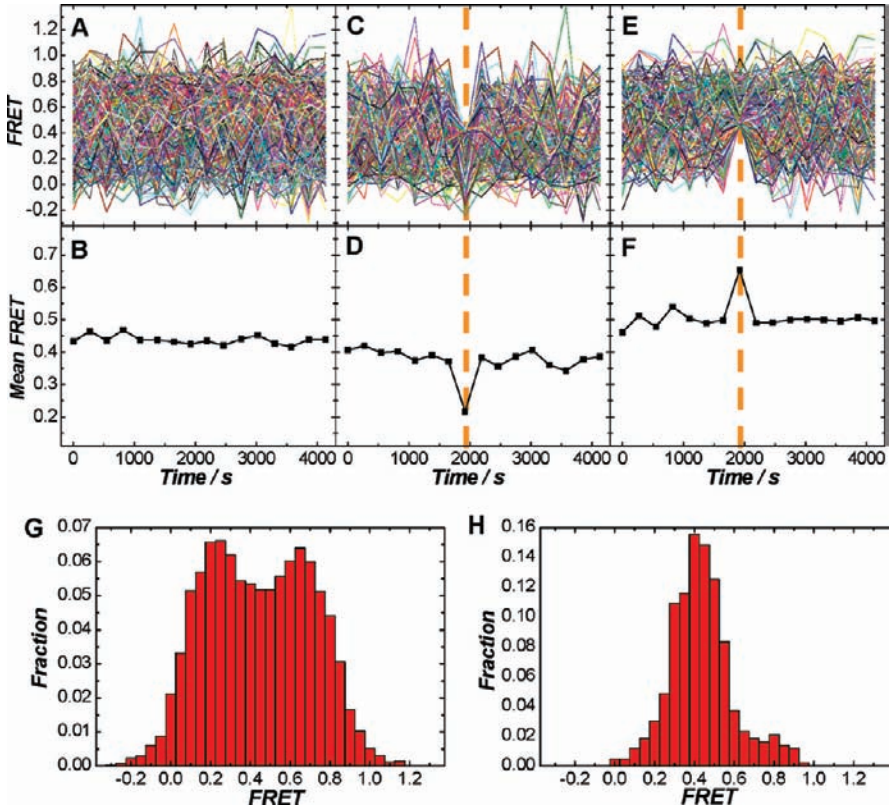

Figure 5. Conformational dynamics of TAR−cTAR DNA duplexes in 800 nM NC studied by SM-FRET measurements using the image scanning mode. Panel A shows the FRET trajectories of 431 molecules, and Panel B shows the corresponding molecularly averaged FRET values. The middle and right columns show the "hole-burning" analysis of the FRET trajectories. Panels C and E show the trajectories of a subpopulation composed of 205 molecules with FRET smaller than 0.45 and a subpopulation composed of 226 molecules with FRET larger than 0.45 at time spot of 1921s (highlighted by dash lines), respectively. The corresponding mean FRET values are shown in Panels D and F. FRET histograms constructed based on (G) all the spectroscopic occurrences of all molecules in this ensemble and (H) the time-averaged FRET values of individual molecules.

competitors, such as Mg^{2+} and yeast *t*RNA (Figure 8), further verifying that NC, as a DNA-bending protein, indeed functions through its binding to dsDNA.

NC Mutants. NC has two functional domains, the basic N-terminal domain and two zinc fingers. The NC's zinc fingers account for NC's ability to locally weaken the base pairs in double-stranded regions of both nucleic acid hairpins[34,37] and duplexes,[8,42] while the positively charged basic domain is crucial to the electrostatic interactions of NC with NAs and the binding cooperativity arising from protein−protein interactions.[3,6] As shown in Figure 9, NC mutants lacking either the basic N-terminal domain or the zinc fingers are unable to effectively bend dsDNA even at micromolar concentration range. The apparent dissociation constants (K_d) for NC binding to dsDNA have been measured to be in the range of 10−100 nM.[43] The NC mutant lacking the N-terminal domain we used is a truncated NC(11−55). Compared to the wild-type NC, NC(11−55) has less positive charge and lacks the N-terminal domain which is responsible for the binding cooperativity and NC's ability to induce NC-NA coaggregates.[6,10,37,44] NC(11−55) has smaller binding affinity to various NAs than wild-type NC; however, the difference in the K_d values of wild-type NC and NC(11−55) to various NAs has been measured to be within 1 order of magnitude.[6,37,44,45] To disrupt the properly folded zinc finger structures, we added 0.5 mM ethylenediaminetetraacetic acid (EDTA) into the NC solutions. EDTA is a chelating ligand that strongly interacts with zinc ions and therefore can be used to effectively extract the zinc ions out of the zinc fingers of NC.[43] Once losing the zinc ions, the zinc fingers can no longer maintain their properly folded structures that are crucial to NC's ability to destabilize the double-stranded regions of NA hairpins or duplexes. However, NC (zinc-free) binds strongly to NAs without significant loss in binding affinity, as NC can fold into various conformations to accommodate productive electrostatic interactions with the NA substrates even in the absence of Zn^{2+} binding.[43,45]

Our SM-FRET results clearly show that both the zinc fingers and N-terminal domain are indispensable for the NC-induced NA bending over nanometer length scale. Although wild-type NC and NC mutant lacking the zinc fingers can both induce effective condensation of the proviral DNA globally,[5,6] our SM-FRET results clearly indicate that the zinc fingers play a crucial role in inducing the local bending of DNA on the nanometer length scale. This furthermore implies that NC's ability to produce sharp bends in duplex NAs is not a simple consequence of binding by a highly positive protein. As a test of this latter hypothesis, we undertook SM-FRET measurements on TAR−

(42) Williams, M. C.; Gorelick, R. J.; Musier-Forsyth, K. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 8614–8619.
(43) 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.; Rein, A.; Rouzina, I.; Williams, M. C. *Nucleic Acids Res.* **2006**, *34*, 593–605.
(44) Godet, J.; de Rocquigny, H.; Raja, C.; Glasser, N.; Ficheux, D.; Darlix, J. L.; Mely, Y. *J. Mol. Biol.* **2006**, *356*, 1180–1192.
(45) Urbaneja, M. A.; Kane, B. P.; Johnson, D. G.; Gorelick, R. J.; Henderson, L. E.; Casas-Finet, J. R. *J. Mol. Biol.* **1999**, *287*, 59–75.

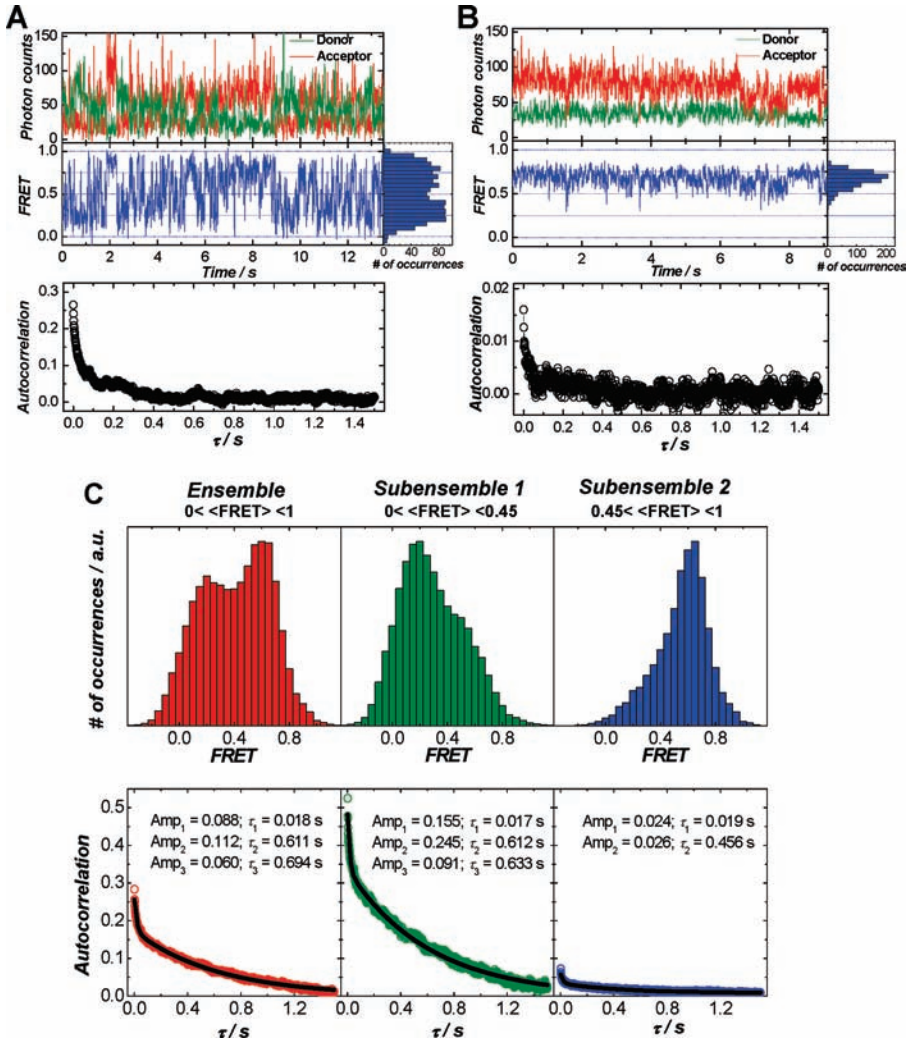

Figure 6. Conformational dynamics of the DNA duplexes in 800 nM NC studied by SM-FRET measurements using the individual trajectory mode: (A and B) Donor, acceptor, FRET trajectories, and FRET autocorrelation collected on two molecules representing two types of kinetics we observed. (C) The ensemble FRET histograms (upper row) and the ensemble FRET autocorrelations (lower row). The left column shows the ensemble FRET histograms and autocorrelation obtained from FRET trajectories of 45 molecules. The middle column shows the subensemble FRET histograms and autocorrelation obtained from subpopulation 1 (21 molecules with time-averaged ⟨FRET⟩ from 0 to 0.45). The right column shows the subensemble FRET histograms and autocorrelation obtained from subpopulation 2 (24 molecules with ⟨FRET⟩ from 0.45 to 1).

cTAR DNA duplexes in the presence of polylysine even at micromolar concentration range (Figure 10). Although polylysine has been widely used to induce condensation of long dsDNAs,[46] no appreciable bending on the nanometer length scale is observed due to the presence of polylysine. This implies that the mechanism by which NC induces DNA duplex condensation is distinct from that of the better known examples of polycation-induced DNA condensation.[46−48]

Hypothetical Origins of NC-Induced Duplex NA Bending. In order to allow the sharp bending of short dsDNA segments to occur at reduced energy-cost, localized deformation of the DNA's double helical structures upon NC binding is indispensable. Rigid materials under bending force typically relieve most of the bending strain at buckled sites, and this applies to the duplexed NA bending as well. Locally melted or kinked regions are expected in sharply bent DNAs.[39,49] Alternatively, intercalation without melting involving the aromatic residues of NC may also contribute to the NC-induced DNA bending. We hypothesize that the ability of NC to make sharp bends in NA duplexes is related to its ability to shift the secondary structure of NAs toward more melted sequences. While the absence of terminal fraying and duplex dissociation indicates that the NC-coated duplexes are substantially in a duplex form rather than melted form, the presence of NC-stabilized non-base-paired "bubbles" that locally disrupt the DNA duplex cannot be ruled out from these experiments. Indeed, NC-induced bubbles have been reported in temperature-induced melting of DNA duplexes,[7] and NC has been known to be able to create locally melted regions within DNA duplexes through the interactions of its zinc fingers with the nucleic acids.[8,42] A bubble could hypothetically function as a kink or hinge for a sharp NA bend. A NC-stabilized bubble located near the midpoint along the length of the duplex might be further stabilized by additional "bridging" NC molecules between the two "arms" of the duplex. This would be consistent with the strong so-called "aggregating" ability of NC,[6,37,44,50] which in this case might favor a bent structure where one or

(46) Behr, J. P. *Acc. Chem. Res.* **1993**, *26*, 274–278.
(47) Tang, M. X.; Szoka, F. C. *Gene Ther.* **1997**, *4*, 823–832.
(48) Bloomfield, V. A. *Biopolymers* **1997**, *44*, 269–282.
(49) Yan, J.; Marko, J. F. *Phys. Rev. Lett.* **2004**, *93*, 108108.
(50) Tsuchihashi, Z.; Brown, P. O. *J. Virol.* **1994**, *68*, 5863–5870.

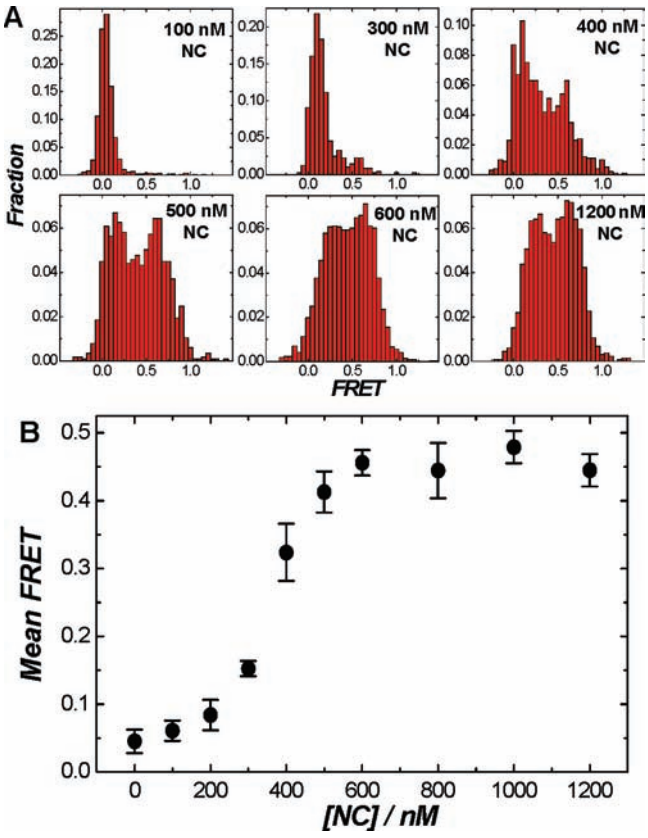

Figure 7. (A) FRET histograms (obtained from the image scanning mode) of the TAR−cTAR DNA duplexes as a function of NC concentration. (B) Molecularly averaged FRET values as a function of NC concentration.

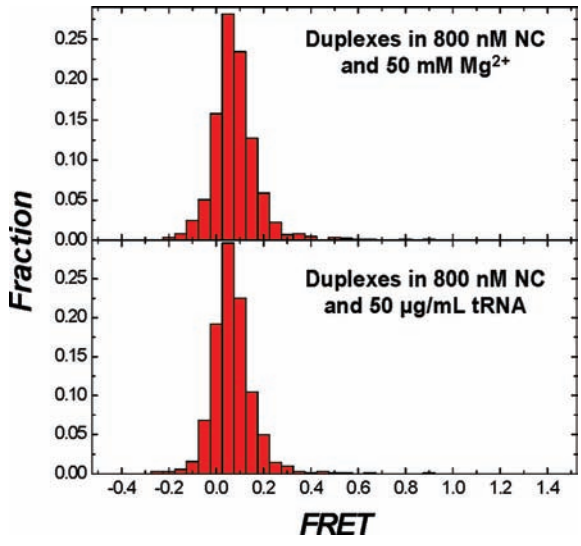

Figure 8. Bending of DNA duplexes by NC in the presence of excessive binding competitors. FRET histograms of the TAR−cTAR DNA duplex molecules in (upper) 800 nM NC and 50 mM Mg^{2+} and (lower) 800 nM NC and 50 μg/mL yeast *t*RNA.

more NC molecules could interact with both arms simultaneously. Selective binding of bridging NC molecules could also lead to asymmetric charge neutralization of the phosphate backbone which has previously been shown to generate a lateral force to induce the spontaneous DNA bending.[51]

(51) Strauss, J. K.; Maher, L. J. *Science* **1994**, *266*, 1829–1834.

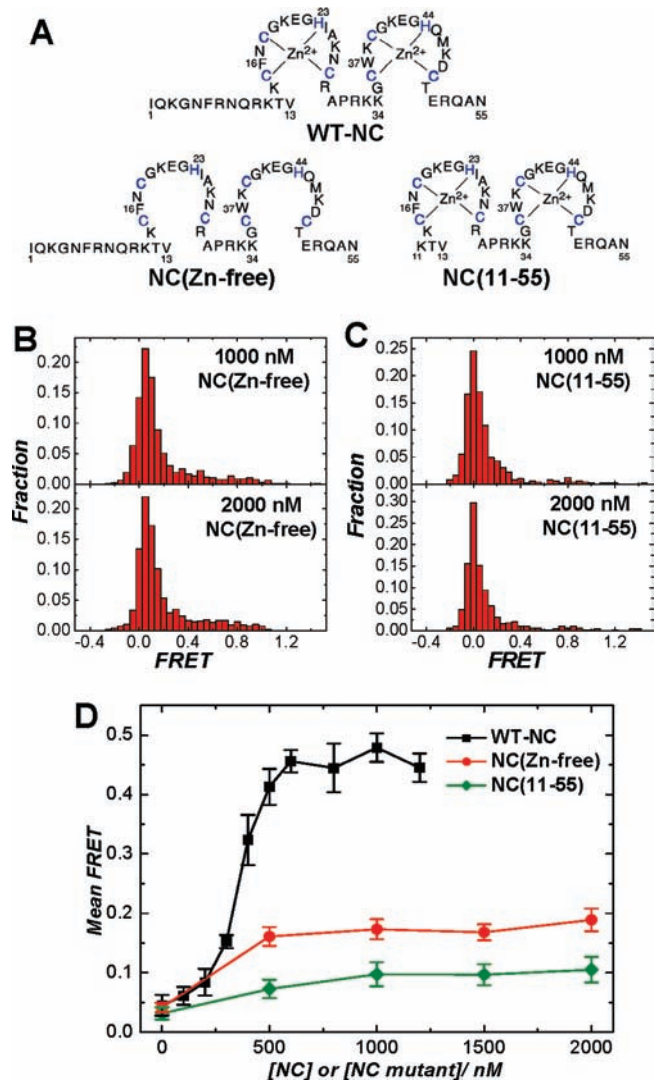

Figure 9. (A) Primary structures of the wild-type NC, NC(Zn-free), and NC(11−55). (B) FRET histograms of the TAR−cTAR DNA duplex molecules in (upper) 1000 nM and (lower) 2000 nM NC(Zn-free). (C) FRET histograms of the DNA duplex molecules in (upper) 1000 nM and (lower) 2000 nM NC(11−55). (D) Mean FRET values as a function of NC or NC mutant concentrations.

Potential Biological Significance. The discovered ability of NC to induce bends in NA duplex on the nanometer length scale in vitro implies several possible biological roles for NC in vivo. It has been reported that HIV-1 NC can induce effective condensation of proviral DNA, which ensures the protection of the newly synthesized proviral DNA during its nuclear import and subsequent integration.[5,6,9] Although the compactness of the proviral DNA coated with NC is comparatively much less than the chromosomal genome compaction in eukaryotes, such NC-induced DNA condensation does provide effective protection of the proviral DNA against cellular nuclease digestion.[6] Although the detailed mechanism has not been fully clarified, the stimulation of the coupled joining of the two ends of the proviral DNA in the host cell genome by NC during the integration step has been clearly verified by in vitro experiments.[9] It is believed that the condensation of dsDNA by NC facilitates the association of the two ends of proviral DNA at precisely spaced locations in an integration complex. It has also been hypothesized that NC may play an indirect role in provirus integration by directing the integrase toward the two ends of

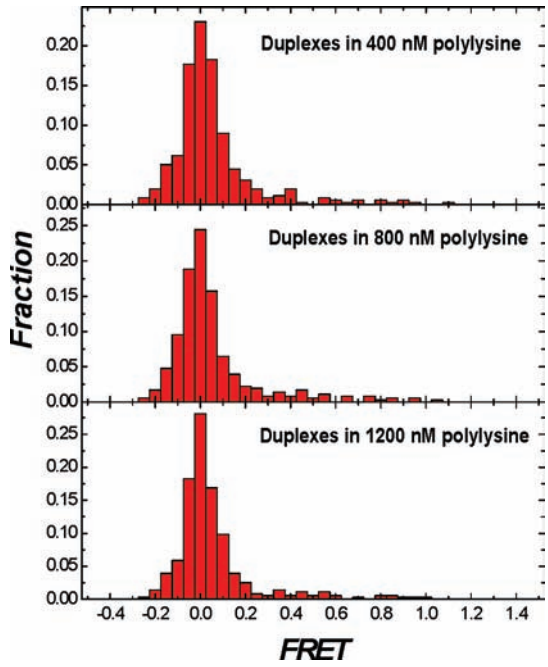

Figure 10. FRET histograms of the TAR−cTAR DNA duplex molecules in (top) 400 nM, (middle) 800 nM, and (bottom) 1200 nM polylysine.

the proviral DNA.[5] Integrase does not specifically recognize DNA ends; however, the fact that NC does not bind efficiently to the DNA ends[5,52] makes it possible for NC to facilitate integrase recognition of the DNA ends by coating the proviral DNA while leaving the DNA ends exposed for integrase cleavage and integration. Although NC-induced proviral DNA condensation has been observed in vitro and its relevant biological roles has been either demonstrated or hypothesized, the molecular level understanding on the mechanisms of this DNA condensation process is not clear yet.

NC-induced localized DNA kinks may give rise to some distinct states of the globally condensed DNA chains and may contribute to the conformational heterogeneity of the nucleoprotein complexes and the kinetic heterogeneity of the DNA condensation processes. It is noteworthy that these local DNA kinks are structurally flexible, as the DNA−NC nucleoprotein complexes show highly dynamic conformational changes over a broad distribution of time scales. Such structural flexibility implies NC's multiple biological functions as a DNA bending protein. It allows NC to reduce the DNA's rigidity globally without forming rigid local structures that might inhibit the interactions of other viral or cellular proteins with the proviral DNA. The flexible local DNA bending also implies a more flexible structure of the condensed proviral DNA compared to that of eukaryotic chromatin. Such flexible, less condensed conformations of the proviral DNA allow a series of viral or cellular proteins to access the proviral DNA for a multitude of biological tasks, such as the transport and integration of proviral DNA. The flexible DNA bending may also allow NC to facilitate the formation of higher order nucleoprotein complexes, such as the preintegration complexes,[53,54] that may require different enzyme-specific local DNA bend angles.

Conclusion

We report the discovery that a short segment of HIV-1 proviral DNA is bent on the nanometer length scale spontaneously at room temperature when the DNA segments are coated with multiple HIV-1 NC protein molecules. Using SM-FRET to probe the NC-induced DNA bending process, we have observed highly heterogeneous conformational dynamics of the nucleoprotein complexes at single-molecular level. In striking contrast to naked DNA in buffer solutions which shows relatively static extended conformations, the DNA bound with NC shows multiple dynamic bent conformations which interconvert over a broad distribution of time scales from milliseconds to minutes. It is proposed that the NC-induced bending of NA involves the formation of NC-stabilized non-base-paired "bubbles" that locally disrupt the NA duplex and act as hinge for a sharp NA bend. It is further envisaged that the NC-stabilized bubble located near the midpoint of the duplex is stabilized by additional "bridging" NC molecules between the two "arms" of the duplex.

Acknowledgment. This work is supported by the Welch Foundation. We thank Karin Musier-Forsyth of The Ohio State University for helpful discussions.

Supporting Information Available: Additional information as noted in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

JA9070046

(52) Sherman, P. A.; Dickson, M. L.; Fyfe, J. A. *J. Virol.* **1992**, *66*, 3593–3601.

(53) Brown, P. O.; Bowerman, B.; Varmus, H. E.; Bishop, J. M. *Cell* **1987**, *49*, 347–356.

(54) Farnet, C. M.; Haseltine, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 4164–4168.