A Mechanism for Plus-Strand Transfer Enhancement by the HIV-1 Nucleocapsid Protein during Reverse Transcription † , ‡

ARTICLE *in* BIOCHEMISTRY · SEPTEMBER 2000

Impact Factor: 3.02 · DOI: 10.1021/bi000841i · Source: PubMed

7 AUTHORS, INCLUDING:

Philip E Johnson
York University
36 PUBLICATIONS **970** CITATIONS
SEE PROFILE

Michael F Summers
University of Maryland, Baltimore County
164 PUBLICATIONS **11,837** CITATIONS
SEE PROFILE

A Mechanism for Plus-Strand Transfer Enhancement by the HIV-1 Nucleocapsid Protein during Reverse Transcription[†,‡]

Philip E. Johnson,[§] Ryan B. Turner,[§,||] Zheng Rong Wu,[§,⊥] Lea Hairston,[§] Jianhui Guo,[#] Judith G. Levin,*[,#] and Michael F. Summers*[,§]

Howard Hughes Medical Institute and Department of Chemistry and Biochemistry, University of Maryland Baltimore County, Baltimore, Maryland 21250, and Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892

Received April 13, 2000; Revised Manuscript Received June 9, 2000

ABSTRACT: The HIV-1 nucleocapsid protein (NC) functions as a nucleic acid chaperone during the plus-strand transfer step in reverse transcription by facilitating annealing of the primer binding site (PBS) sequence in the short plus-strand strong-stop DNA fragment [(+) SSDNA] to a complementary site located near the 3′ end of the minus-strand DNA [(−) PBS DNA]. To investigate the mechanism by which NC performs this function, we have prepared an 18-nucleotide (−) PBS DNA for nuclear magnetic resonance (NMR) based structural and NC binding studies. The (−) PBS DNA forms a stable hairpin ($T_m \sim 42 \pm 5$ °C) that contains a five-residue loop and a bulged thymine in a guanosine−cytosine-rich stem. Addition of substoichiometric amounts of NC results in significant broadening and reductions in NMR signal intensities of the Watson−Crick base-paired imino protons and a reduction by 20 °C in the upper temperature at which the imino proton signals are detectable, consistent with destabilization of the structure. The results suggest that inefficient annealing in the absence of NC may be due to the intrinsic stability of an internal (−) PBS DNA hairpin and that NC facilitates strand transfer by destabilizing the hairpin and exposing stem nucleotides for base pairing with the PBS sequence in (+) SSDNA.

The genomes of all retroviruses encode a Gag polyprotein that is produced in the infected cell during the late stages of the infectious cycle (*1*). Approximately 1500 copies of Gag associate at the cell membrane and bud to form an immature viral particle (*2*). Concomitant with budding, the Gag polyproteins are cleaved by the viral protease to produce the viral structural proteins. In the case of human immunodeficiency virus type 1 (HIV-1),[1] these include matrix (MA, p17), capsid (CA, p24), p2, nucleocapsid (NC, p7), p1, and p6 proteins (*3*−*5*), which rearrange during a process called maturation to form the infectious virus particle (*1*).

The nucleocapsid proteins play important roles in several stages of the retroviral replication cycle. Prior to proteolytic cleavage, the NC domains of the Gag polyproteins participate in the recognition and packaging of the viral genome (see ref *6* and references therein) and also appear to be important for virus particle formation (*7*, *8*) and placement of the tRNA reverse transcription primer [$tRNA_{Lys3}$ for HIV-1 (*1*)] at the primer binding site (PBS) in genomic RNA (*9*−*11*). After proteolytic cleavage, the mature nucleocapsid proteins form a ribonucleoprotein complex that stabilizes the genomic RNA dimer in the mature virus (*1*, *12*; see also refs *13*−*16*).

After viral entry, a linear, double-stranded DNA copy of the viral RNA genome is synthesized by the virion-associated reverse transcriptase (RT) enzyme via a process that is also dependent on NC. RT-dependent DNA synthesis proceeds via a multistep process that involves two strand transfer events (*17*−*20*) (Figure 1). Previous studies have shown that NC facilitates the HIV-1 minus- and plus-strand DNA transfer steps (*21*−*32*) (see Figure 1), which are necessary for elongation of the two DNA strands and for synthesis of the long terminal repeat present at each end of proviral DNA (*19*). During plus-strand transfer, the initial, short product made during initiation of plus-strand synthesis, termed plus-strand strong-stop DNA [(+) SSDNA], is translocated to the 3′ end of minus-strand viral DNA (*19*). The actual transfer

[†] This work was supported by NIH Grant GM42561 to M.F.S. and in part by funds from the NIH Intramural AIDS Targeted Antiviral Program awarded to J.G.L. P.E.J. was supported by a postdoctoral fellowship from the Natural Sciences and Engineering Research Council of Canada, R.B.T. was supported by Meyerhoff, MARC U*STAR, and Barry Goldwater undergraduate fellowships, and L.H. was supported by NIH MBRS-IMSD Grant R25GM55036.

[‡] The coordinates for an ensemble of 20 structures of the (−) PBS DNA have been deposited in the Protein Data Bank (1EN1).

* To whom correspondence should be addressed. M.F.S.: phone, (410) 455-2880; fax, (410) 455-1174; e-mail, summers@hhmi.umbc.edu. J.G.L.: phone, (301) 496-1970; fax, (301) 496-0243; e-mail, judith_levin@nih.gov.

[§] University of Maryland Baltimore County.

[||] Current address: Harvard Medical School, Cambridge, MA.

[⊥] Current address: Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892.

[#] National Institutes of Health.

[1] Abbreviations: Ade, adenosine; Cyt, cytidine; Gua, guanosine; HIV-1, human immunodeficiency virus type 1; HMQC, heteronuclear multiple-quantum coherence; NC, nucleocapsid protein; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; PBS, primer binding site; RT, reverse transcriptase; (+) SSDNA, plus-strand strong-stop DNA; (−) PBS DNA, primer binding site located near the 3′ end of minus-strand DNA; Thy, thymidine; TOCSY, total correlation spectroscopy; T_m, melting temperature.

10.1021/bi000841i CCC: $19.00
Published on Web 07/15/2000

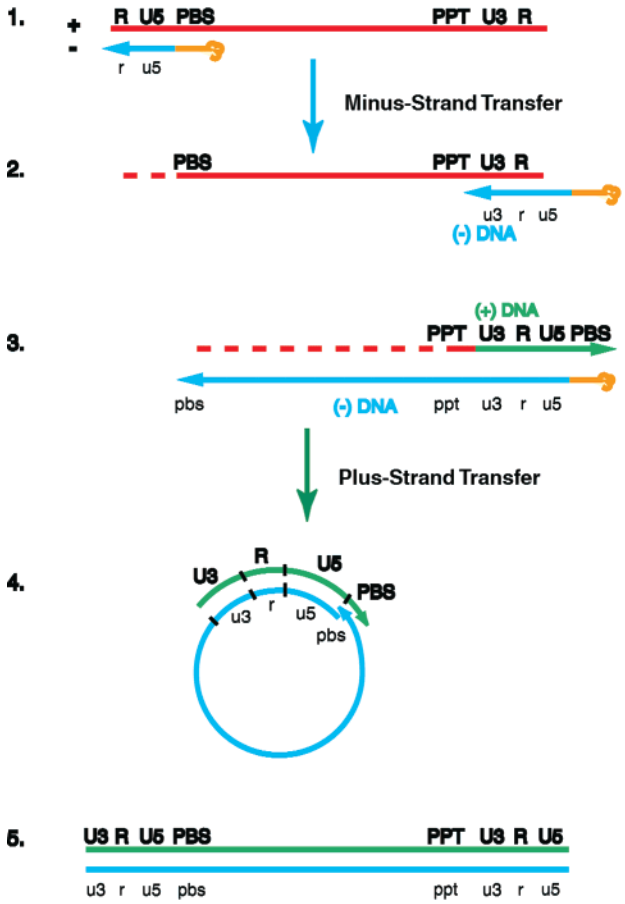

FIGURE 1: Schematic diagram illustrating the steps in reverse transcription. (Reaction 1) Reverse transcription is initiated by a cellular tRNA primer. The 3′ 18 nucleotides of the tRNA are annealed to the primer binding site (PBS) near the 5′ end of viral RNA. Reverse transcriptase catalyzes extension of the primer; this results in synthesis of (−) SSDNA, which contains a copy of the repeat (R) region and unique 5′ RNA sequences (U5). (Reaction 2) As (−) SSDNA is being synthesized, the RNase H activity of RT degrades the 5′ viral RNA sequences that have been reverse transcribed. Transfer of (−) SSDNA (with the tRNA primer still attached) to the 3′ end of viral RNA (minus-strand transfer) is facilitated by annealing of the complementary R regions in (−) SSDNA and viral RNA. (Reaction 3) As minus-strand DNA is elongated, genomic RNA sequences are degraded by RNase H, and plus-strand DNA synthesis is initiated by the polypurine tract (PPT) primer. RT copies the minus-strand unique 3′ sequences (u3), r, and u5, as well as the 18 nucleotides at the 3′ end of the tRNA primer; this reconstitutes the PBS. RNase H removes the PPT and tRNA primers (not shown). (Reaction 4) Transfer of (+) SSDNA to the 3′ end of minus-strand DNA (plus-strand transfer) is facilitated by annealing of the complementary PBS sequences at the 3′ termini of (+) SSDNA and minus-strand DNA. The two DNA strands circularize during the transfer event, and displacement synthesis follows. (Reaction 5) Elongation of plus- and minus-strand DNAs results in formation of a linear double-stranded DNA with a long terminal repeat (LTR) at each end. Minus-strand and plus-strand DNA sequences are in lower and upper case letters, respectively. Viral RNA is shown in red, (−) strand DNA in blue, (+) strand DNA in green, and tRNA in orange. The two strand transfer events are indicated with a blue arrow (minus-strand transfer) and a green arrow (plus-strand transfer). The dashed red lines represent RNase H cleavage of viral RNA sequences. The diagram is adapted from Telesnitsky and Goff (*78*).

event requires annealing of the complementary 18-nucleotide PBS sequences at the 3′ ends of the plus- and minus-strand DNA molecules (Figure 1).

The mechanism by which NC facilitates strand transfer during reverse transcription is not known but may be generally related to the protein's intrinsic nucleic acid chaperone activity. Thus, the NC protein has been shown to be capable of unwinding tRNA (*33*, *34*), facilitating the renaturation of complementary DNA strands (*35*) and DNA strand exchange (*36*), promoting annealing in minus-strand (*24*, *27*, *28*, *32*, *79*) and plus-strand (*29*, *79*) transfer, inducing maturation of the dimeric RNA genome (*10*, *13*, *15*, *16*), reducing RT pausing during strand elongation (*37*−*39*), and promoting annealing of the tRNA primer to the PBS sequence (*10*, *40*−*48*). It has been suggested that NC promotes these activities because it can induce the transient unpairing of the bases within a chain, to allow alternative, more stable combinations of base pairs to form (*36*, *49*).

Interestingly, several groups have reported that a number of other nucleic acid binding proteins are unable to substitute for NC in the tRNA unfolding (*33*), tRNA−PBS annealing (*32*, *50*), and minus-strand transfer (*26*) reactions that occur during reverse transcription. In contrast, however, NC interacts with a variety of viral and nonviral nucleic acid sequences and can catalyze the hybridization of complementary RNA−RNA, RNA−DNA, and DNA−DNA sequences (*24*, *27*−*29*, *32*, *35*, *50*, *79*).

The structure of the HIV-1 NC protein has been determined both free in solution (*34*, *51*−*53*) and bound specifically to stem−loop recognition elements of the Ψ-RNA packaging signal (*54*, *80*), revealing substantial changes in the conformation and dynamics of the protein and RNA upon binding. HIV-1 NC contains two conserved "zinc knuckle" domains (Cys-X_2-Cys-X_4-His-X_4-Cys; X = variable amino acid) that bind specifically to exposed guanosine nucleobases in the RNA loops, as well as a stretch of basic residues in the N-terminal tail that interact with the RNA phosphodiester groups. These studies provided insights into the mechanism of sequence-specific NC−RNA interactions, as well as a rationale for the high conservation of more than 50% of the HIV-1 NC amino acids. The structures of a short DNA fragment (dACGCC) bound to the isolated N-terminal zinc knuckle of HIV-1 NC (*55*) and to the intact NC protein (*56*) also revealed an affinity of the zinc knuckle domains for guanosine.

Although information regarding the structural determinants of genome recognition by NC is gradually being revealed, insights into the molecular mechanisms of NC-dependent strand transfer and chaperone activities remain poor, at best. In the present study, NMR methods have been employed to investigate the structure of the 18-nucleotide minus-strand (−) PBS DNA and the nature of its interactions with NC. The (−) PBS DNA forms a stable stem−loop structure, and addition of NC leads to NMR spectral changes consistent with a reduction in stability, or an increase in conformational lability, of the DNA stem−loop. These studies suggest that NC facilitates the plus-strand transfer event by destabilizing an otherwise stable (−) PBS DNA internal stem−loop structure.

MATERIALS AND METHODS

Protein and DNA Sample Preparation. (−) PBS DNA used for NMR experiments [5′-d(Gua-Thy-Cyt-Cyt-Cyt-Thy-Gua-Thy-Thy-Cyt-Gua-Gua-Gua-Cyt-Gua-Cyt-Cyt-Ade)] was

obtained from Oligos Etc. Inc. (Wilsonville, OR); the DNA oligonucleotide was purified by the company using an HPLC chromatographic procedure that yields a product with greater than 90% purity. The lyophilized DNA for NMR analysis was resuspended in either H_2O or ^2H_2O (99.9%, Isotech). The concentration of DNA used for the NMR experiments was approximately 2 mM. Recombinant NC protein samples [corresponding to the HIV-1_{NL4-3} isolate (*57*)] were prepared and purified as described (*53*, *54*). The zinc content (2 equiv of Zn per NC protein) was established by 1D ^{1}H NMR and mass spectrometry.

Electrophoresis Measurements. The oligomerization state of the (−) PBS DNA oligonucleotide was analyzed by gel electrophoresis under nondenaturing conditions (12% polyacrylamide). A 2 μL sample of the (−) PBS DNA (~0.2 mM) was mixed with 2 μL of 50% glycerol and then loaded onto the gel. TBE buffer (*58*) was used as the running solution. Samples of the 19-nucleotide RNA hairpin SL2, the 20-nucleotide RNA hairpin SL3, and a 50-nucleotide construct containing both the SL2 and SL3 sequences from the HIV-1 encapsidation site were electrophoresed in neighboring lanes to provide size/molecular weight comparisons.

NMR Spectroscopy. NMR experiments were performed using Bruker DMX 600 MHz and GE Omega PSG 600 MHz NMR spectrometers equipped with triple or single *z*-axis gradient capabilities. ^{1}H NMR assignments were obtained from a combination of two-dimensional homonuclear clean-TOCSY (*59*) and NOESY (mixing time; $\tau_m = 200$ ms) (*60*) experiments recorded in 99.9% ^2H_2O at 25 °C. A NOESY experiment (mixing time; $\tau_m = 200$ ms) with the DNA in 90% H_2O/10% ^2H_2O at 5 °C was also collected, with water suppression achieved through the use of the WATERGATE sequence (*61*). A natural abundance ^{1}H−^{13}C HMQC experiment was used to unambiguously identify the adenosine H2 signal (*62*). One-dimensional ^{1}H NMR experiments of the (−) PBS DNA in 90% H_2O/10% ^2H_2O were collected at 5 °C in the absence and presence of HIV-1 NC (DNA:NC ratios of 1:0.1, 1:0.5, and 1:0.9) using a binomial pulse train for water suppression (*63*). NMR spectra were processed and analyzed with the NMRPipe (*64*) and NMRView (*65*) software packages, respectively.

DNA Melting Experiments. An estimate of the melting temperature of the (−) PBS DNA hairpin was obtained by monitoring the NMR chemical shift changes of the nonexchangeable protons as a function of temperature. One-dimensional ^{1}H spectra of the DNA in ^2H_2O were obtained at temperatures of 5−65 °C in 5 °C steps. The chemical shifts of well-resolved signals were plotted against temperature, with the midpoint of the resulting sigmoidal curve used to determine the melting point. Plots were obtained for three well-resolved signals of the DNA stem. Additional melting experiments were performed over the same temperature range with the sample in 90% H_2O/10% ^2H_2O to monitor changes in the imino region of the DNA. Melting experiments were conducted in both the presence and absence of NC protein.

Structure Calculations. Structure calculations were performed using DYANA (*66*). Distance restraints of 1.8−2.7, 1.8−3.3, and 1.8−5.0 Å were employed for strong, medium, and weak intensity cross-peaks, respectively, observed in the 2D NOESY spectrum. An additional 0.5 Å was added for NOEs involving methyl protons. Hydrogen bond restraints for each G-C Watson−Crick base pair employed lower and upper limit restraints of 1.80 and 1.89 Å, respectively, for the GuaH1−CytN3, GuaH21−CytO2, and GuaO6−CytH41 distances, and lower and upper limit restraints of 2.90 and 3.00 Å, respectively, for the GuaN1−CytN3, GuaN2−CytO2, and GuaO6−CytN4 distances. Loose torsion angle restraints were used for the DNA phosphodiester groups (α between −115° and −15°, β between 125° and 225°, γ between 20° and 120°, ϵ between −220° and −120°, ζ between −150° and −50°) in order to maintain reasonable phosphodiester torsion angles and allow sampling of both A and B helical conformations (*67*). Structures were analyzed and figures produced with the Midas (*68*), Molscript (*69*), and Setor (*70*) software packages.

RESULTS AND DISCUSSION

The (−) PBS DNA Oligonucleotide Is Monomeric. Nondenaturing gel electrophoresis experiments were performed to determine if the (−) PBS DNA oligonucleotide exists as a monomer or duplex under the conditions employed. Included on the gel as size standards were samples of the RNA stem loops SL2 (19 nucleotides), SL3 (20 nucleotides), and SL2-3 (50 nucleotides) from the HIV-1 Ψ-RNA packaging signal, which were shown previously to form monomeric stem−loop structures (*54*, *71*, *81*). The 18-nucleotide (−) PBS DNA migrated slightly faster than the 20-nucleotide SL3 and 19-nucleotide SL2 RNA and much faster than the 50-nucleotide SL2-3 RNA, indicating that the (−) PBS DNA oligonucleotide is also monomeric (data not shown). In addition, the ^{1}H NMR spectrum of the (−) PBS DNA oligonucleotide exhibited a single set of signals over the concentration range 0.2−2.0 mM, consistent with the presence of a single monomeric species. These data indicate that the (−) PBS DNA is monomeric in solution and does not form a duplex under the conditions employed.

^{1}H NMR Assignment of (−) PBS DNA. Complete assignments for the nonexchangeable base hydrogen and deoxyribose H1′, H2′, and H2″ signals, and nearly complete assignments of the remaining deoxyribose protons, were obtained by analysis of standard connectivities observed in the 2D NOESY and TOCSY spectra (*72*). A selected portion of a 2D NOESY spectrum showing sequential connectivities between H6/H8 and H1′ is shown in Figure 2. Assignments of all the exchangeable protons involved in the four Watson−Crick base pairs except for Cyt14 H41 and H42 were also made. The absence of signals for the Cyt14 NH_2 group likely results from fast proton exchange occurring at the Gua1-Cyt14 base pair, consistent with its location at the base of the stem and adjacent to the bulged Thy2 nucleotide.

Structure of the (−) PBS DNA Stem Loop. An ensemble of 20 structures of the (−) PBS DNA oligonucleotide with a mean target function of 1.17 ($\pm$0.63) $\times$ 10^{-3} Å^2 was generated with DYANA using 72 NOE-derived upper limit distance restraints and hydrogen bond restraints for four Gua-Cyt base pairs. Residues of the stem exhibit high convergence (superposition of all heavy atoms in residues Gua1, Cyt3−Cyt5, and Gua11−Cyt14 affords pairwise RMS deviations of 1.65 $\pm$ 0.29 Å), but residues Thy2, Thy6, Gua7, Thy8, Thy9, and Cyt10 are not as well defined relative to the rest of the structure (Table 1 and Figure 3).

Residues Gua1, Cyt3, Cyt4, and Cyt5 form Watson−Crick base pairs with residues Cyt14, Gua13, Gua12, and Gua11,

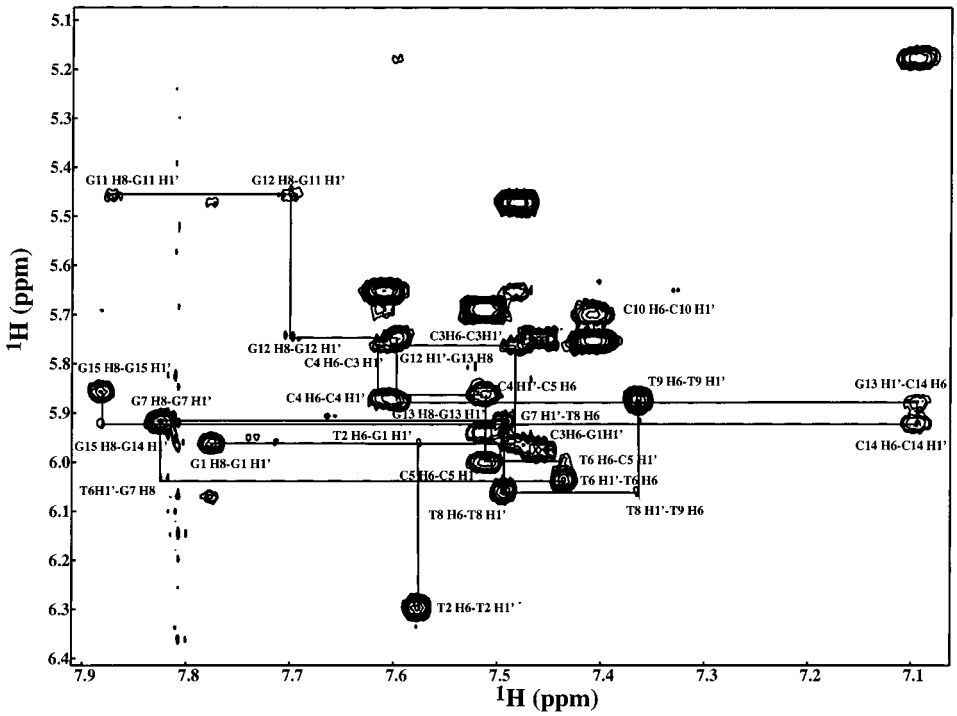

FIGURE 2: An expanded region of the two-dimensional homonuclear NOESY spectrum of the (−) PBS DNA oligonucleotide in 99.9% ^2H_2O. This portion of the spectrum shows the NOE walk connecting sequential and intranucleotide H1′ and H6/H8 protons. The lines drawn show the connectivities used to assign the DNA; for clarity only selected peaks are labeled.

Table 1: Distance Restraints and Structural Statistics for the Final Ensemble of 20 Structures for the (−) PBS DNA Hairpin

(A) distance restraints[a]	
intraresidue	14
interresidue	58
hydrogen bond	48
total NMR-derived restraints	120
average restraints per refined residue[b]	13.7
(B) distance violations	
mean total penalty (Å^2)	$(1.17 \pm 0.63) \times 10^{-3}$
minimum total penalty (Å^2)	0.63×10^{-3}
maximum total penalty (Å^2)	2.50×10^{-3}
(C) pairwise RMSD (Å) rel to mean atom positions	
base-paired nucleotides[c] (all heavy atoms)	1.65 ± 0.29
all structured stem−loop nucleotides[d] (all heavy atoms)	2.60 ± 0.54

[a] Intraresidue and interresidue restraints employed only upper limit restraints; hydrogen bond restraints employed both upper and lower limit restraints. [b] Nucleotides included are Gua1−Cyt14. [c] Nucleotides included are Gua1, Cyt3−Cyt5, and Gua11−Cyt14. [d] This does not include the bulged Thy2 and includes Gua1 and Cyt3−Cyt14.

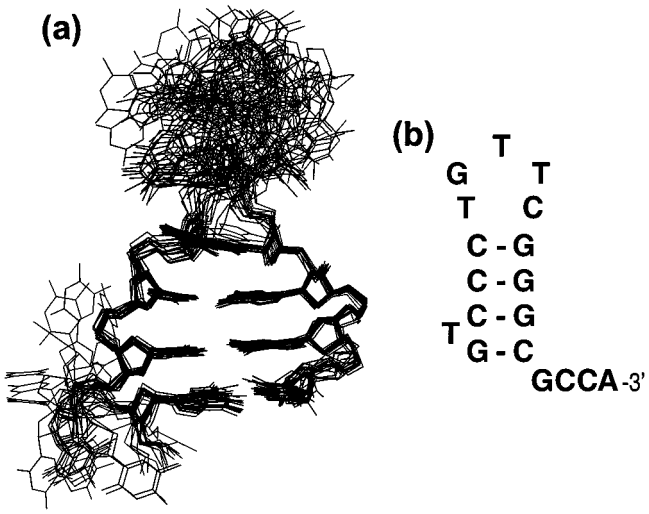

FIGURE 3: (a) Superposition of the ensemble of 20 refined structures of the (−) PBS DNA. The bulged nucleotide Thy2 in the stem as well as the nucleotides in the pentaloop are clearly disordered in this ensemble of structures. This figure was generated by the superposition of the heavy atoms of the structured residues in the stem of the molecule, Gua1, Cyt3−Cyt5, and Gua11−Cyt14. All hydrogen atoms and oxygen atoms of the phosphodiester backbone are omitted for clarity. (b) Schematic diagram of the secondary structure of the (−) PBS DNA.

respectively, and exhibit an overall B-helical structure. NOE interactions between Gua1 H8 and Cyt3 H5 and H6, and between Gua1 H2″ and Cyt3 H6, provide evidence that Gua1 and Cyt3 are stacked in a B-helical geometry and that Thy2 is not stacked in the DNA stem. In fact, no internucleotide NOEs were observed for the Thy2 methyl group, providing further evidence that the Thy2 base is looped out of the stem and forms a bulge. The position of Thy2 relative to the rest of the stem is poorly defined in the structures due to a paucity of internucleotide NOEs, and it is possible that Thy2 is conformationally mobile.

Nucleotides Thy6, Gua7, Thy8, Thy9, and Cyt10 form a partially ordered pentaloop (Figures 3 and 4). Although it is not apparent in Figure 3, a self-consistent subset of NOEs indicates that portions of the pentaloop contain at least transient structural elements. NOE interactions between Thy6 CH3 and Cyt5 H6, and between Thy6 CH3 and Cyt5 H2′ and H2″(along with the expected Thy6 H6−Cyt 5 H1′ spin-diffusion cross-peak, Figure 2), indicate that the first nucleotide in the loop, Thy6, stacks on Cyt5 (Figure 4). The methyl group of Thy8 exhibits NOEs with protons on the ribose ring of Gua7 and Gua7 H8, and these NOEs position the Thy8 nucleobase near the stem. The methyl group of

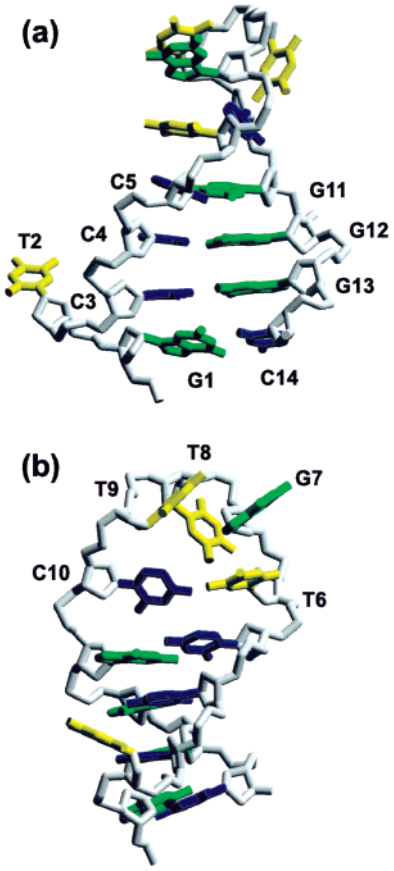

FIGURE 4: Two views of the structure of the (−) PBS DNA stem loop having the lowest target function. The top view (a) shows the bulged Thy2 nucleotide as well as the bases forming the stem. The bottom view (b) shows the position of the nucleotides in the pentaloop. Guanosine residues are shown in green, thymines in yellow, and cytosines in blue.

Thy9 exhibits numerous cross-peaks with the ribose protons of Thy8, in addition to cross-peaks with the Gua7 H8, H2′, and H1′ protons and with Thy8 H6. Also, Thy9 H6 exhibits cross-peaks with the Gua7 H1′ proton. These data are consistent with the pentaloop structure shown in Figure 4, in which the nucleobase of Thy9 points across the loop in the direction of Thy6.

The final nucleotide of the pentaloop, Cyt10, exhibits NOEs from its H6 proton to the Thy8 CH3 and H2′ protons, and from its H5 proton to the Thy9 H2′. There is no evidence that Cyt10 stacks with Gua11. The observed NOEs position Cyt10 on the inside of the loop, pointing toward Thy6 as shown in Figure 4.

Melting Stability of the (−) PBS DNA Oligonucleotide in the Absence and Presence of NC. One-dimensional NMR spectra of the (−) PBS DNA oligonucleotide were obtained as a function of temperature to determine the stability of the hairpin structure. Sigmoidal temperature-dependent chemical shifts observed for several nucleobase protons revealed a melting midpoint temperature of 42 ± 5 °C (data not shown).

We were unable to obtain similar thermal melting data for the DNA in the presence of NC due to severe line broadening and signal overlap with the NC protein resonances. However, good quality NMR spectra of the Watson−Crick imino protons were obtained in the presence and absence of NC and were used to investigate the stability of the stem. As shown in Figure 5, relatively sharp signals

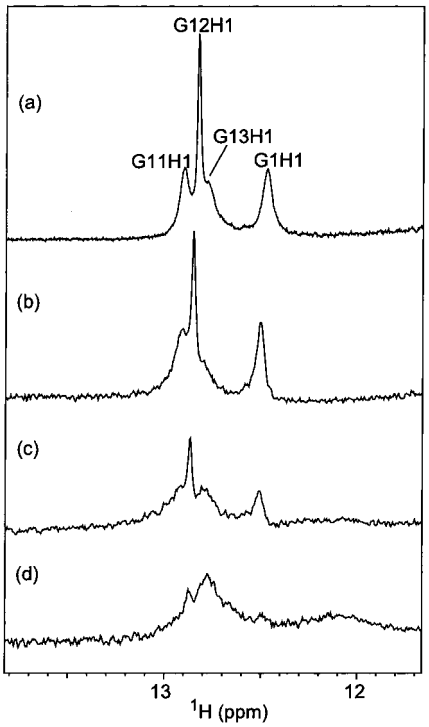

FIGURE 5: Destabilization of the (−) PBS DNA stem−loop upon HIV-1 nucleocapsid protein binding. Shown here are the one-dimensional ^{1}H NMR spectra of the (−) PBS DNA focusing on the imino signals resulting from the four Watson−Crick base pairs. These spectra are of the free DNA (a) and those upon the addition of NC protein at DNA to protein molar ratios of 1:0.1 (b), 1:0.5 (c), and 1:0.9 (d). The large increase in line widths and decrease in intensity of the sharp signals in the 12−13 ppm region in (a) indicate that binding of the DNA by NC destabilizes the base pairing of the DNA structure. These data provide direct evidence that a function of NC is to destabilize local nucleic acid structure that could inhibit duplex formation with (+) PBS DNA during reverse transcription.

between 12 and 13 ppm were observed for Gua imino protons involved in Gua-Cyt base pairing in the absence of NC. Successive titrations of NC to the (−) PBS DNA oligonucleotide resulted in a pronounced broadening of these signals (Figure 5), as well as to a general broadening of the resolved signals of the nonexchangeable protons. The increased line broadening shown in Figure 5 is indicative of the destabilization of the Watson−Crick base pairs in the stem and clearly demonstrates that NC binding disrupts the structure of the (−) PBS stem. Other regions of the NMR spectrum provide direct evidence that the N-terminal zinc knuckle of NC participates directly in binding. In particular, the appearance of a signal at 0.1 ppm is consistent with signals assigned previously to Ile24 CH3$^\gamma$ in protein and peptide complexes with DNA and RNA oligonucleotides that contain a guanosine (*54*, *55*). Only one set of NMR signals was observed for the (−) PBS DNA oligonucleotide upon titration with NC, indicating that binding occurs in the fast-exchange regime of the NMR chemical shift time scale (millisecond).

The ^{1}H NMR spectra of the Watson−Crick imino proton signals were also monitored as a function of temperature in the presence and absence of NC. The maximum temperature at which the imino protons of the free (−) PBS DNA oligonucleotide were visible was 40 °C. This is only slightly lower than the midpoint of the thermal melting transition

(42 ± 5 °C) and is due to the fact that the imino protons exchange rapidly with water protons when the Gua-Cyt base pairs are broken and the imino protons are exposed to solvent. Most importantly, the maximum temperature at which the imino proton signals were observable for the DNA in the presence of 0.9 equiv of NC was only 20 °C, which reflects a 20 °C reduction compared to the spectra obtained for the free DNA. As indicated above, signal loss results from rapid exchange of the imino protons with water protons. These data therefore clearly demonstrate that NC binding destabilizes the stem, resulting in an increased exposure of the stem imino protons to water.

Biological Relevance. The reverse transcription of the HIV-1 genome in vivo depends on the presence of the native NC protein (*73*−*75*). As described above, in vitro studies indicate that the HIV-1 NC protein functions during reverse transcription by increasing both the rate and extent of minus- and plus-strand transfer (Figure 1). The NMR studies presented here focus on the plus-strand transfer event and demonstrate that the (−) PBS DNA is capable of forming a stable stem−loop structure. Thus, it is possible that inefficient strand transfer in the absence of NC may be due to the trapping of the (−) PBS sequence in a relatively stable stem−loop structure, which prevents its efficient base pairing with the complementary PBS sequence in (+) SSDNA. Our studies demonstrate that NC is capable of destabilizing the (−) PBS DNA stem−loop, resulting in the exposure of nucleotides that were sequestered in the stem. Thus, it appears that NC may facilitate strand transfer by destabilizing the stem and exposing the stem nucleotides for base pairing with the complementary (+) PBS sequence in (+) SSDNA.

Because binding occurs in the fast-exchange regime of the NMR chemical shift time scale, it was not possible to obtain high-resolution structural information for the complex or complexes that form upon titration of the (−) PBS DNA oligonucleotide with NC. Thus, it is not clear if NC binds to one or more sites on the (−) PBS DNA stem loop or exists in one or more rapidly interconverting conformations upon binding. However, based on NMR spectral comparisons with data obtained previously for NC protein and zinc knuckle peptide binding to other DNA and RNA oligonucleotides, it is clear that the NC zinc knuckle domains are direct participants in binding. Previous studies have shown that HIV-1 NC preferentially binds guanosine nucleotides via a hydrophobic cleft in a zinc knuckle (*55*). In the structure of the HIV-1 NC−SL3 complex, the two exposed guanosines in the tetraloop are bound by the two zinc knuckles of a single NC protein (*54*). In both cases, the binding of exposed guanosine bases to the N-terminal zinc knuckle leads to dramatic upfield chemical shifts of the Phe16 and Ile24 proton signals, with the Ile24 $CH3^{\gamma}$ signal shifting to ca. 0.1 ppm. Addition of NC to the (−) PBS DNA oligonucleotide results in a similar 0.1 ppm signal, which strongly suggests that the N-terminal zinc knuckle of NC interacts with one or more guanosines upon complex formation.

At least two binding models that involve interactions between the N-terminal zinc knuckle of NC and guanosine nucleobases could explain the observed destabilization of the hairpin structure. In one model, the NC protein initially binds the single guanosine (Gua7) in the pentaloop. The exocyclic ring of this guanosine is solvent exposed in a similar manner as seen for the guanosines in the loop of the unbound SL3 RNA (*71*); it is precisely these exposed loop guanosines in the SL3 loop that bind to the zinc knuckles of the NC protein upon tight, 1:1 complex formation (*54*). Note that Gua7 is sandwiched between two thymine residues, and this TGT sequence has been previously shown to serve as an optimal binding site for the N-terminal zinc knuckle (*55*) and also for the intact NC protein (*76*, *77*). Since there is only one guanosine in the loop of the (−) PBS DNA oligonucleotide, the second zinc knuckle domain might bind to a stem guanosine (possibly Gua11, located closest to the loop), which would disrupt the Watson−Crick base pair at the top of the stem (Cyt5-Gua11) and possibly increase the accessibility of other guanosines for interactions with additional NC molecules or with the complementary PBS sequence in (+) SSDNA.

An alternate binding model also assumes the interaction of HIV-1 NC with guanosine nucleobases, but in this model the NC protein binds to guanosines during the time in which the stem opens or "breathes" while it is in equilibrium with an open or duplex conformation. The sensitivity of the imino proton signals of the terminal Watson−Crick base pairs to chemical exchange with solvent protons is indicative of "fraying", which is a common phenomenon, and could easily facilitate the binding of the NC zinc knuckles. Thus, during the brief time a guanosine is unpaired, it is may be susceptible to being captured by the zinc knuckles of NC. The adjacent guanosines in the Gua-Cyt base pairs would then be more susceptible to unfolding. In both models, the binding of NC to the (−) PBS DNA would make the nucleobases in the stem more accessible to water, giving rise to the observed increase in imino proton exchange rates, and would thus also make the stem nucleotides more accessible for forming base pairs with the (+) PBS sequence in (+) SSDNA. The basic residues of NC associated with the N-terminal tail and the linker peptide that connects the zinc knuckles are probably also involved in destabilization of the stem, and in this regard, future studies of (−) PBS DNA interactions with peptides comprising only the basic segments or the isolated zinc knuckle domains may shed light on their relative contributions to DNA binding and stem destabilization.

It is noteworthy that the binding of HIV-1 NC to the SL3 RNA hairpin at 1:1 stoichiometry does not lead to the destabilization of the stem of this molecule (*54*). The SL3 tetraloop contains a high-affinity NC binding site with two exposed guanosines, and when NC binds to SL3, the two zinc knuckles bind tightly and specifically to the exposed guanosine nucleobases of the tetraloop. However, when the concentration of NC exceeds that of RNA, the ^{1}H NMR signals of the RNA become broadened in a manner similar to that observed in the present study (unpublished results). Thus, NC appears to also be capable of destabilizing the SL3 structure once the high-affinity binding site has been saturated and the NC:SL3 ratio exceeds 1:1.

The present studies have demonstrated that the 18-nucleotide (−) PBS DNA forms a hairpin structure that contains a structured pentaloop and a bulged Thy2 nucleotide in the stem. The stem is relatively stable, with a thermal melting temperature of approximately 42 °C. Addition of substoichiometric amounts of NC results in a destabilization of the structure, as evidenced by significant broadening and reductions in NMR signal intensities of the Watson−Crick

base-paired imino protons. The studies suggest that inefficient annealing in the absence of NC (*29*, *79*) is due to the presence of a stable (−) PBS DNA internal hairpin and that NC facilitates annealing with the (+) PBS sequence in (+) SSDNA by destabilizing the hairpin structure. Although high-resolution structural studies of the complexes formed upon titration of the (−) PBS DNA oligonucleotide with NC were precluded by fast-exchange binding that probably involves multiple species, future structural studies of complexes with mutant DNA sequences may allow the testing of the possible mechanisms of NC-dependent chaperone and strand transfer activities.

ACKNOWLEDGMENT

We are grateful to Gaya Amarasinghe (HHMI, UMBC) for helpful discussions, to Rossi Gitti and Robert Edwards (HHMI, UMBC) for technical assistance, and to Louis Henderson (SAIC-Frederick, NCI-FCRDC) and Alan Rein (NCI, NCI-FCRDC) for useful comments on the manuscript. We thank Klara Post (LMG, NICHD, NIH) for help in generating Figure 1.

REFERENCES

1. Coffin, J. M., Hughes, S. H., and Varmus, H. E. (1997) *Retroviruses*, Cold Spring Harbor Laboratory Press, Plainview, NY.
2. Vogt, V. M., and Simon, M. N. (1999) *J. Virol. 73*, 7050−7055.
3. Mervis, R. J., Ahmad, N., Lillehoj, E. P., Raum, M. G., Salazar, F. H. R., Chan, H. W., and Venkatesan, S. (1988) *J. Virol. 62*, 3993−4002.
4. Henderson, L. E., Bowers, M. A., Sowder, R. C., II, Serabyn, S. A., Johnson, D. G., Bess, J. W., Jr., Arthur, L. O., Bryant, D. K., and Fenselau, C. (1992) *J. Virol. 66*, 1856−1865.
5. Vogt, V. M. (1997) in *Retroviruses* (Coffin, J. M., Hughes, S. H., and Varmus, H. E., Eds.) pp 27−69, Cold Spring Harbor Laboratory Press, Plainview, NY.
6. Berkowitz, R., Fisher, J., and Goff, S. P. (1996) *Curr. Top. Microbiol. Immunol. 214*, 177−218.
7. Franke, E. K., Yuan, H. E. H., Bossolt, K. L., Goff, S. P., and Luban, J. (1994) *J. Virol. 68*, 5300−5305.
8. Swanstrom, R., and Wills, J. W. (1997) in *Retroviruses* (Coffin, J. M., Hughes, S. H., and Varmus, H. E., Eds.) pp 263−334, Cold Spring Harbor Press, Plainview, NY.
9. Huang, Y., Khorchid, A., Gabor, J., Wang, J., Li, X., Darlix, J.-L., Wainberg, M. A., and Kleiman, L. (1998) *J. Virol. 72*, 3907−3915.
10. Feng, Y.-X., Campbell, S., Harvin, D., Ehresmann, B., Ehresmann, C., and Rein, A. (1999) *J. Virol. 73*, 4251−4256.
11. Fu, W., Ortiz-Conde, B. A., Gorelick, R. J., Hughes, S. H., and Rein, A. (1997) *J. Virol. 71*, 6940−6946.
12. Bolognesi, D. P., Montelaro, R. C., Frank, H., and Schäfer, W. (1978) *Science 199*, 183−186.
13. Fu, W., Gorelick, R. J., and Rein, A. (1994) *J. Virol. 68*, 5013−5018.
14. Tanchou, V., Gabus, C., Rogemond, V., and Darlix, J.-L. (1995) *J. Mol. Biol. 252*, 563−571.
15. Muriaux, D., de Rocquigny, H., Roques, B. P., and Paoletti, J. (1996) *J. Biol. Chem. 271*, 33686−33692.
16. Feng, Y.-X., Copeland, T. D., Henderson, L. E., Gorelick, R. J., Bosche, W. J., Levin, J. G., and Rein, A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 7577−7581.
17. Gilboa, E., Mitra, S. W., Goff, S., and Baltimore, D. (1979) *Cell 18*, 93−100.
18. Arts, E. J., and Wainberg, M. A. (1996) *Adv. Virus Res. 46*, 97−163.
19. Telesnitsky, A., and Goff, S. P. (1997) in *Retroviruses* (Coffin, J. M., Hughes, S. H., and Varmus, H. E., Eds.) pp 121−160, Cold Spring Harbor Laboratory Press, Plainview, NY.
20. Whitcomb, J. M., and Hughes, S. H. (1992) *Annu. Rev. Cell Biol. 8*, 275−306.
21. Darlix, J.-L., Vincent, C., Gabus, H., de Rocquigny, H., and Roques, B. (1993) *C. R. Acad. Sci. 361*, 763−771.
22. Allain, B., Lapadat-Tapolsky, M., Berlioz, C., and Darlix, J.-L. (1994) *EMBO J. 13*, 973−981.
23. Peliska, J. A., Balasubramanian, S., Giedroc, D. P., and Benkovic, S. J. (1994) *Biochemistry 33*, 13817−13823.
24. You, J. C., and McHenry, C. S. (1994) *J. Biol. Chem. 269*, 31491−31495.
25. Guo, J., Henderson, L. E., Bess, J., Kane, B., and Levin, J. G. (1997) *J. Virol. 71*, 5178−5188.
26. Kim, J. K., Palaniappan, C., Wu, W., Fay, P. J., and Bambara, R. A. (1997) *J. Biol. Chem. 272*, 16769−16777.
27. Davis, W. R., Gabbara, S., Hupe, D., and Peliska, J. A. (1998) *Biochemistry 37*, 14213−14221.
28. Guo, J., Wu, J., Bess, J., Henderson, L. E., and Levin, J. G. (1998) *J. Virol. 72*, 6716−6724.
29. Wu, T., Guo, J., Bess, J., Henderson, L. E., and Levin, J. G. (1999) *J. Virol. 73*, 4794−4805.
30. Raja, A., and DeStefano, J. J. (1999) *Biochemistry 38*, 5178−5184.
31. Auxilien, S., Keith, G., Le Grice, S. F., and Darlix, J.-L. (1999) *J. Biol. Chem. 274*, 4412−4420.
32. Lapadat-Tapolsky, M., Pernelle, C., Borrie, C., and Darlix, J. L. (1995) *Nucleic Acids Res. 23*, 2434−2441.
33. Khan, R., and Giedroc, D. P. (1992) *J. Biol. Chem. 267*, 6689−6695.
34. Summers, M. F., Henderson, L. E., Chance, M. R., Bess, J. W. J., South, T. L., Blake, P. R., Sagi, I., Perez-Alvarado, G., Sowder, R. C. I., Hare, D. R., and Arthur, L. O. (1992) *Protein Sci. 1*, 563−574.
35. Dib-Hajj, F., Khan, R., and Giedroc, D. P. (1993) *Protein Sci. 2*, 231−243.
36. Tsuchihashi, Z., and Brown, P. O. (1994) *J. Virol.*, 5863−5870.
37. Wu, W., Henderson, L. E., Copeland, T. D., Gorelick, R. J., Bosche, W. J., Rein, A., and Levin, J. G. (1996) *J. Virol. 70*, 7132−7142.
38. Ji, Z., Klarmann, G. J., and Preston, B. D. (1996) *Biochemistry 35*, 132−143.
39. Drummond, J. E., Mounts, P., Gorelick, R. J., Casas-Finet, J. R., Bosche, W. J., Henderson, L. E., Waters, D. J., and Arthur, L. O. (1997) *AIDS Res. Hum. Retroviruses 13*, 533−543.
40. Prats, A.-C., Sarih, L., Gabus, C., Litvak, S., Keith, G., and Darlix, J.-L. (1988) *EMBO J. 7*, 1777−1783.
41. Darlix, J.-L., Gabus, C., Nugeyre, M.-T., Clavel, F., and Barre-Sinoussi, F. (1990) *J. Mol. Biol. 216*, 689−699.
42. de Rocquigny, H., Gabus, C., Vincent, A., Fournié-Zaluski, M.-C., Roques, B., and Darlix, J.-L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6472−6476.
43. Barat, C., Schatz, O., Le Grice, S., and Darlix, J.-L. (1993) *J. Mol. Biol. 231*, 185−190.
44. Li, X., Quan, Y., Arts, E., Li, Z., Preston, B. D., DeRocquigny, H., Roques, B. P., Darlix, J.-L., Kleiman, L., Parniak, M. A., and Wainberg, M. A. (1996) *J. Virol. 70*, 4996−5004.
45. Huang, Y., Khorchid, A., Wang, J., Parniak, M. A., Darlix, J. L., Wainberg, M. A., and Kleiman, L. (1997) *J. Virol. 71*, 4378−4384.
46. Remy, E., DeRocquigny, H., Petitjean, P., Muriaux, D., Theilleux, V., Paoletti, J., and Roques, B. P. (1998) *J. Biol. Chem. 273*, 4819−4822.
47. Chan, B., and Musier-Forsyth, K. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 13530−13535.
48. Chan, B., Weidemaier, K., Yip, W.-T., Barbara, P. F., and Musier-Forsyth, K. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 459−464.
49. Rein, A., Henderson, L. E., and Levin, J. G. (1998) *Trends Biochem. Sci. 23*, 297−301.
50. Darlix, J.-L., Lapadat-Tapolsky, M., de Rocquigny, H., and Roques, B. P. (1995) *J. Mol. Biol. 254*, 523−537.
51. Morellet, N., Jullian, N., de Rocquigny, H., Maigret, B., Darlix, J.-L., and Roques, B. P. (1992) *EMBO J. 11*, 3059−3065.

52. Omichinski, J. G., Clore, G. M., Sakaguchi, K., Appella, E., and Gronenborn, A. M. (1991) *Fed. Eur. Biochem. Soc. Lett. 292*, 25−30.
53. Lee, B. M., De Guzman, R. N., Turner, B. G., Tjandra, N., and Summers, M. F. (1998) *J. Mol. Biol. 279*, 633−649.
54. De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N., and Summers, M. F. (1998) *Science 279*, 384−388.
55. South, T. L., and Summers, M. F. (1993) *Protein Sci. 2*, 3−19.
56. Morellet, N., Demene, H., Teilleux, V., Huynh-Dinh, T., de Rocquigny, H., Fournie-Zaluski, M.-C., and Roques, B. P. (1998) *J. Mol. Biol. 283*, 419−434.
57. Adachi, A., Gendelman, H. E., Koenig, S., Folks, T., Willey, R., Rabson, A., and Martin, M. A. (1986) *J. Virol. 59*, 284−291.
58. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
59. Griesinger, C., Otting, G., Wüthrich, K., and Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870−7872.
60. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546−4553.
61. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
62. Bax, A., and Subramanian, S. (1986) *J. Magn. Reson. 67*, 565−569.
63. Plateau, P., and Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310−7311.
64. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
65. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR 4*, 603−614.
66. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
67. Wutke, D. S., Foster, M. P., Case, D. A., Gottesfeld, J. M., and Wright, P. E. (1997) *J. Mol. Biol. 273*, 183−206.
68. Ferrin, T. E., Huang, C. C., Jarvis, L. E., and Langridge, R. (1988) *J. Mol. Graphics 6*, 13−27.
69. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
70. Evans, S. V. (1993) *J. Mol. Graphics 11*, 134−138.
71. Pappalardo, L., Kerwood, D. J., Pelczer, I., and Borer, P. N. (1998) *J. Mol. Biol. 282*, 801−818.
72. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
73. Tanchou, V., Decimo, D., Péchoux, C., Lener, D., Rogemond, V., Berthoux, L., Ottmann, M., and Darlix, J.-L. (1998) *J. Virol. 72*, 4442−4447.
74. Gorelick, R. J., Gagliardi, T. D., Bosche, W. J., Wiltrout, T. A., Coren, L. V., Chabot, D. J., Lifson, J. D., Henderson, L. E., and Arthur, L. O. (1999) *Virology 256*, 92−104.
75. Gorelick, R. J., Fu, W., Gagliardi, T. D., Bosche, W. J., Rein, A., Henderson, L. E., and Arthur, L. O. (1999) *J. Virol. 73*, 8185−8195.
76. Fisher, R. J., Rein, A., Fivash, M., Urbaneja, M. A., Casas-Finet, J. R., Medaglia, M., and Henderson, L. E. (1998) *J. Virol. 72*, 1902−1909.
77. Vuilleumier, C., Bombarda, E., Morellet, N., Gerard, D., Roques, B. P., and Mely, Y. (1999) *Biochemistry 38*, 16816−16825.
78. Telesnitsky, A., and Goff, S. P. (1993) in *Reverse Transcriptase* (Skalka, A. M., and Goff, S. P., Eds.) pp 49−83, Cold Spring Harbor Press, Cold Spring Harbor, NY.
79. Guo, J., Wu, T., Anderson, J., Kane, B. F., Johnson, D. G., Gorelick, R. J., Henderson, L. E., and Levin, J. G. (2000) *J. Virol.* (in press).
80. Amarasinghe, G. K., De Guzman, R. N., Turner, R. B., Chancellor, K., Wu, Z.-R., and Summers, M. F. (2000) *J. Mol. Biol.* (in press).
81. Amarasinghe, G. K., De Guzman, R. N., Turner, R. B., and Summers, M. F. (2000) *J. Mol. Biol. 299*, 145−156.

BI000841I