Evidence for the Differential Effects of Nucleocapsid Protein on Strand Transfer in Various Regions of the HIV Genome*

Received for publication, November 18, 2002, and in revised form, February 14, 2003
Published, JBC Papers in Press, February 20, 2003, DOI 10.1074/jbc.M211701200

Suchitra S. Derebail, Megan J. Heath, and Jeffrey J. DeStefano‡
From the Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, Maryland 20742

An in vitro strand transfer assay that mimicked recombinational events occurring during reverse transcription in HIV-1 was used to assess the role of nucleocapsid protein (NC) in strand transfer. Strand transfer in highly structured nucleic acid species from the U3 3' long terminal repeats, gag-pol frasetrnshame region, and Rev response element were strongly enhanced by NC. In contrast, weakly structured templates from the env and pol-env regions transferred well without NC and showed lower enhancement. The lack of strong polymerase pause sites in the latter regions demonstrated that non-pause driven mechanisms could also promote transfer. Assays conducted using NC zinc finger mutants supported a differential role for the two fingers in strand transfer with finger 1 (N-terminal) being more important on highly structured RNAs. Overall this report suggests a role for structural intricacies of RNA templates in determining the extent of influence of NC on recombination and illustrates that strand transfer may occur by several different mechanisms depending on the structural nature of the RNA.

All retroviruses are characterized by reverse transcription, which involves the conversion of viral RNA into double-stranded DNA that becomes integrated into the host cellular chromosome. This process is initiated from a nucleoprotein complex in the virion core that is primarily composed of the dimerized diploid RNA genome coated with nucleocapsid protein (NC)1 (1, 2). In addition the viral enzymes such as reverse transcriptase (RT), integrase (IN), and protease (PR) are present along with host tRNA molecules. RT carries out DNA synthesis during reverse transcription whereas NC serves as a co-factor in this process and others (see below). Nucleocapsid protein is a highly basic, positively charged protein, comprised of 55 amino acid residues. It is derived by proteolysis of the gag and gag-pol precursor polyprotein (3, 4). All retroviral NC proteins contain either one or two conserved 14-residue invariant motifs, Cys-X2-Cys-X2-His-X2-Cys, where X denotes variable amino acids. They are also known as the zinc finger motifs, because they bind zinc ions strongly (5–7). The two zinc fingers of HIV are in close spatial proximity (8–10) and have similar structures (11) but are not functionally equivalent (12, 13). NC exhibits a multitude of essential functions in the life cycle of HIV (14–18) that make it an ideal target for drug therapy (19–21) and vaccine development (22). It has been shown to unwind (17, 23) and anneal the tRNA1-RNA primer to the RNA genome (24–26). NC also participates in recognition and packaging of the viral genome (13, 27–29), in the maturation of genomic RNA dimer (30–32), and possibly in integration of the double-stranded DNA into the host genome (33). In addition, NC serves as a nucleic acid chaperone (14, 34) capable of destabilizing secondary structures and enhancing the annealing of complementary nucleic acids.

Previous studies have demonstrated that one or more of this array of functions enables NC to enhance strand transfer (a process that leads to recombination) during HIV replication (35–37). Strand transfer involves the switching of DNA being synthesized on one RNA template (referred to as “donor”) to homologous regions on the same or on a second RNA template (referred to as “acceptor”) where the synthesis continues. There are two obligatory strand transfers that occur at the termini of the viral RNA (38–40). These are integral steps without which the viral replication cannot proceed to completion. In addition to these vital events, the virus can also undergo strand transfers at internal regions of the RNA template that can potentially occur at any position along the genome and during the syntheses of both minus and plus DNA strands (41–43). These are important steps that aid in generating genetic diversity in the viral population (39, 40, 44, 45) and allow some viruses to evade host immune response and drug therapy. They also may increase the probability of successful DNA synthesis by providing a salvage pathway for broken or damaged genomes (38).

NC has been shown to enhance the two obligatory transfers (35, 37, 46, 47), as well as internal strand transfer events (15, 36, 48, 49). The mechanism by which NC protein promotes these processes is not yet completely understood but presumably stems from a combination of the nucleic acid annealing and helix destabilization activities of NC (36, 37, 50–53). You and McHenry (37) have shown that NC stimulates the annealing of the HIV repeat region sequences by 3000-fold and consequently accelerates strand transfer of terminally redundant sequences. Tsujihashi and Brown (51) have shown that NC promotes strand exchange between double-stranded and single-stranded molecules. Others (15) have shown that NC facilitates strand exchange between donor and acceptor RNAs using model retroviral replication intermediates. A number of other nucleic acid-binding proteins have been shown to facilitate melting and annealing reactions in a similar fashion. Escherichia coli single-stranded DNA-binding (SSB) causes a reduction in the electrostatic repulsion of negatively charged DNA, melts out intramolecular structures, and hence enhances renaturation (54). RNA-binding proteins like heterogeneous

* This work was supported by NIGMS, National Institutes of Health Grant GM51140. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Cell Biology and Molecular Genetics, University of Maryland, Bldg. 231, College Park, MD 20742. Tel.: 301-405-5449; Fax: 301-314-9489; E-mail: jd146@umail.umd.edu.

1 The abbreviations used are: NC, nucleocapsid protein; RT, reverse transcriptase; HIV, human immunodeficiency virus; LTR, long terminal repeats; ssDNA, minus strand strong stop DNA; IN, integrase; PR, protease; RRE, Rev response element; HIV-1, HIV, type 1.
nuclear ribonucleoprotein A1 also accelerate annealing and strand exchange (55).

Though a majority of earlier reports on strand transfer have elaborated upon the activity of NC in highly structured RNA templates, relatively little is known about the effects on RNA templates with weak or no structure. In this report we set out to address two specific questions. (i) Presuming that strand transfer occurred all along the genome (41), does NC exhibit a consistent or a differential influence on strand transfer along the length of the viral genome? For example, in regions with strong secondary structures, NC may be required to unwind these structures to facilitate association of the nascent DNA and acceptor RNA, whereas in regions relatively devoid of stable structures, NC may exert only a minimal influence on strand transfer. (ii) Do the zinc fingers of NC contribute equally or differentially toward strand transfer on various regions of the genome? This question stems from the apparent functional non-equivalence of the two HIV NC zinc fingers noted above. To understand these aspects of NC, strand transfers in five different regions of the HIV genome were analyzed; three of these regions were highly structured whereas two had relatively weak structures. Two of the highly structured regions, the gag-pol frameshift region and Rev response element (RRE), are important HIV regulatory sites. Results demonstrated that even in the absence of NC a relatively high rate of strand transfer occurred in regions with little structure compared with more structured regions. However, NC markedly stimulated the latter in comparison to the former. To evaluate the contributions made by the zinc fingers, the activity levels of mutant NC proteins containing duplicate copies of finger 1 (N-terminal zinc finger) or finger 2 (C-terminal zinc finger) on high or low structure RNA was tested. Results were consistent with finger 1 possessing a larger portion of the protein’s helix-destabilizing activity. Overall, this report establishes the differential effect of NC on strand transfer on various regions of the HIV genome.

**EXPERIMENTAL PROCEDURES**

Plasmid pNL4–3 obtained from the NIH AIDS Research and Reference Reagent Program contains a complete copy of the HIV-1 genome derived from strains NY5 and LAV (56). PCR primers were obtained from Integrated DNA Technologies, Inc. Recombinant HIV-RT was graciously provided by Genetics Institute (Cambridge, MA). This enzyme has a specific activity of about 40,000 units/mg (one unit of RT is defined as the amount required to incorporate 1 nmol of dTTP into nucleic acid product in 10 min at 37 °C using oligo(dT)-polyr(A) as primer-template). The enzyme contained very low levels of single-stranded nuclelease activity, which was found to be inhibited by including 5 mM AMP in the assays (57). At this concentration, the AMP did not affect the polymerase or RNase H activity of RT. Aliquots of HIV-RT were stored frozen at −70 °C, and a fresh aliquot was used for each experiment. HIV NC clone was a generous gift from Dr. Charles McHenry (University of Colorado). NC was purified according to the protocol described (58). The purity of the protein was evaluated using Coomassie Blue staining of 17.5% SDS–PAGE gels (59). Quantification was by absorbance at 280 nm using an extinction coefficient of 8350 cm−1 M−1. Aliquots of NC were stored frozen at −70 °C, and a fresh aliquot was used for each experiment. NC finger mutants 1.1 and 2.2 were a gift from Dr. Robert Gorelick (SAIC, Frederick, MD). These proteins were expressed and purified as described by Bushman and co-workers (60) and quantified by amino acid analysis on a Beckman System 6300 amino acid analyzer (Beckman Coulter, Inc., Fullerton, CA). The polymerase was from Eppendorf. SPS RNA polymerase, DNase I-RNase-free, and RNase-DNase-free were from Roche Diagnostics. RNase inhibitor was from Promega. T4 polynucleotide kinase was obtained from New England Biolabs. Proteinase K was obtained from Eastman Kodak Co. Radiolabeled compounds were as follows:

**TABLE I**

| Region          | Base numbers | Primer sequences |
|-----------------|--------------|------------------|
| U3 3′/LTR donor | 9096–9265    | a. 5′ gattagggacactaatagtatataaagagacagatcct 3′ |
| U3 3′/LTR acceptor | 9074–9245        | b. 5′ tetctattttggctcetetc 3′ |
| PolVif donor    | 3441–3610    | a. 5′ gattagggacactaatagtatatactggaaagctgtaactc 3′ |
| PolVif acceptor | 3419–3590    | b. 5′ tctcctcttgctgctc 3′ |
| GagPol donor    | 2031–2200    | a. 5′ gattagggacactaatagtatataaagagacagatcct 3′ |
| GagPol acceptor | 2009–2180    | b. 5′ tatttctgcggctgcagcctag 3′ |
| Env donor       | 7101–7270    | a. 5′ gattagggacactaatagtatatacctggagaaagctgtaactc 3′ |
| Env acceptor    | 7079–7250    | b. 5′ ctgctgctgctgctgc 3′ |
| RRE donor       | 7823–7992    | a. 5′ gattagggacactaatagtatataaagagacagatcct 3′ |
| RRE acceptor    | 7801–7972    | b. 5′ gatctctccagacagacagacagcctag 3′ |

* Refers to sequence numbering of plasmid pNL4–3 that contains an almost complete copy of the HIV-1 genome derived from strains NY5 and LAV.

* The primer sequences used to prime the pNL4–3 plasmid. The bases in boldface are the SP6 promoter sequences, and those in italics are non-retroviral sequences that were added to prevent end transfers. All the “a” primers are in 3′ orientation, and the “b” primers are in 3′ orientation.

**FIG. 1.** Schematic representation of the strand transfer assay. Donor and acceptor RNA templates are labeled along with their corresponding total lengths. The 150-nucleotide (nt) boxed region enclosing the two templates is the region of homology or the transfer zone. The asterisk represents the 5′-labeled primer that is complementary to only the donor template. The broken lines represent the full-length DNA that is synthesized on the donor, and the dotted lines represent the transfer DNA that has undergone a strand transfer event to the acceptor. The lengths of full-length donor-directed DNA and transfer DNA products are indicated at the bottom of the figure.
Fig. 2. Predicted secondary structures of highly and weakly structured donor RNA templates. RNAdraw and Zuker’s RNA folding programs (see “Results”) were used to predict the secondary structures of donor RNA templates (at a temperature of 37 °C, using program default settings) derived from three regions of the HIV genome. The structures show 155 bases of the donor templates excluding the 20 bases that bind
[γ-32P]ATP was obtained from Amersham Biosciences, Sephadex G-25 spin columns were from Amika Corp., and all other chemicals were from Sigma or Fisher Scientific.

PCR Amplification of DNA Substrates—Five different sets (four primers per set) of PCR primers were specifically designed to yield donor and acceptor RNAs that shared homology over a 150-base region and amplified DNA from five different areas of the pNL4–3 plasmid (see Table I). An SP6 promoter sequence was included on some of the primers to allow transcription of the DNA by SP6 RNA polymerase. PCR reactions were performed according to the enzyme manufacturer’s protocol using the provided buffer. One hundred pmol of each primer was used. The following cycling parameters were used: 35 cycles of

![FIG. 3. Time course strand transfer assays of highly structured RNA substrates.](http://www.jbc.org/content/early/2018/04/11/jbc.M117.794280/F3.large.jpg)

A, the U3 3’LTR donor template is shown (bases 9086–9265, ΔG = −49.30 kcal/mol). 65 marks the position of a major pause site at a U residue (enclosed in the box) that is 65 nucleotides from the 3’ end of the donor template (note that the 20 nucleotides binding to the primer are not shown). #1 and #2 denote the boundaries of the 24-nucleotide potential hotspot (see “Results”). Pause sites within this region have not been mapped. B, GagPol donor template is shown (bases 2031–2200, ΔG = −45.30 kcal/mol). 77 marks the position of a major pause site at an A residue (enclosed in the box). SS refers to the “slippery site” heptamer sequence (enclosed in the box). C, RRE donor template is shown (bases 7823–7992, ΔG = −59.40 kcal/mol). 72 marks the position of a major pause site at an A residue (enclosed in the box). D, PolVif template is shown (bases 3441–3610, ΔG = −19.3 kcal/mol). E, Env template is shown (bases 7101–7270, ΔG = −15 kcal/mol).
denaturation, annealing, and extension at temperatures of 94, 50, and 72 °C, respectively, for 1 min each, one cycle of extension at 72 °C for 5 min. The PCR products were purified on 8% native polyacrylamide gels (29:1) (acrylamide:bisacrylamide) and used to prepare RNA as described below.

**Preparation of RNA Substrates**—Run-off transcription (performed according to the enzyme manufacturer’s protocol) was conducted using ~5 μg of the purified PCR DNAs and SP6 RNA polymerase enzyme to generate donor RNA transcripts of 175 nucleotides and acceptor RNA transcripts of 177 nucleotides. The transcription reactions were treated with 0.4 units/μl of DNase I-RNase-free enzyme for 15 min to digest away the template DNA. Then they were extracted with phenol:chloroform:isoamyl alcohol (25:24:1) and precipitated with ethanol. The RNA pellets were resuspended in 50 μl of RNase-free water, loaded onto hydrated Sephadex G-25 spin columns, and processed according to the manufacturer’s directions. The amount of recovered RNAs was determined spectrophotometrically from optical density.

**RNA-DNA Hybridization**—DNA primers that bound specifically to the donor RNA transcripts were 32P-labeled at the 5′ end with T4 polynucleotide kinase according to the manufacturer’s protocol. Each of the five donor RNAs was hybridized to a complementary labeled primer. Following reagents at the indicated final concentrations were also included in the reaction mixtures: 50 mM Tris-HCl (pH 8.0), 1 mM dithiothreitol, 80 mM KCl, 0.1 mM EDTA (pH 8.0), 6 mM MgCl2 (53), 100 μM dNTPs, 5 mM AMP (pH 7.0), 25 μM ZnCl2, and 0.4 units/μl RNase inhibitor. Reactions were started by the addition of 8 μl of HIV-RT at a final concentration of 2.5 units/μl. The following reagents at the indicated final concentrations were also included in the reaction mixtures: 50 mM Tris-HCl (pH 8.0), 1 mM dithiothreitol, 80 mM KCl, 0.1 mM EDTA (pH 8.0), 6 mM MgCl2 (53) 100 μM dNTPs, 5 mM AMP (pH 7.0), 25 μM ZnCl2, and 0.4 units/μl RNase inhibitor. Reactions were allowed to incubate for time points of 0, 2, 4, 8, 16, 32, and 64 min at 37 °C. At these time points, a 6-μl aliquot of each reaction was terminated by mixing with 4 μl of a solution containing 25 mM EDTA (pH 8.0), 2.5 ng of RNase-DNase-free enzyme and allowed to digest for 20 min at 37 °C. Two μl of proteinase K at 2 mg/ml in 1.25% SDS, 15 mM EDTA (pH 8.0), and 10 mM Tris (pH 8.0) was then added to the above mixture, which was placed at 65 °C for 1 h. Finally, 12 μl of 2× formamide dye (90% formamide, 10 mM EDTA (pH 8.0), 0.1% xylene cyanol, 0.1% bromphenol blue) was added to the mixture, and the samples were resolved on an 8% denaturing polyacrylamide gel containing 7 M urea. Similar reactions were conducted during the analysis of mutant NC proteins (1.1 and 2.2). In the NC titration experiments the total reaction volume was reduced to 12.5 μl, the amount of NC in the reactions was varied as indicated in Fig. 7, and the reactions were allowed to proceed for 32 min. Extended DNA products were quantified by phosphorimager analysis using a GS-525 phosphorimager from Bio-Rad.

**Gel Electrophoresis**—Denaturing 8% polyacrylamide gels (19:1) (acrylamide:bisacrylamide), containing 7 M urea, and native 8% polyacrylamide gels (29:1) (acrylamide:bisacrylamide) were prepared and subjected to electrophoresis as described (59).

**Fig. 4. Graph of efficiency of strand transfer versus time for the highly structured RNA substrates.** Graphs were derived from the experiments shown in Fig. 3. The percent transfer efficiency was defined as Transfer DNA products (T)/Transfer + Full-length donor-directed products (F), times 100 ((T/(T+F)) × 100). The templates used are indicated above each graph. The solid circles represent time course reactions conducted in the absence of nucleocapsid protein (−NC) and the presence of acceptor RNA (+Accep). The open circles represent time course reactions conducted in the presence of nucleocapsid protein (+NC) and acceptor RNA (+Accep).
RESULTS

Strand Transfer Assay—The general approach used to test for strand transfer in the different regions of the HIV genome is depicted in Fig. 1. This assay is designed to simulate internal strand transfer events occurring during minus strand DNA synthesis. The donor (template on which DNA synthesis initiates) and acceptor (template to which DNAs initiating on the donor can potentially transfer) represent the two strands of viral genomic RNA in the virion. DNA synthesis is initiated from a 5’ end-labeled DNA primer that was specifically designed to bind only to the 3’ end of the donor RNA. Strand transfer can occur over the transfer zone, which is the region of homology between the donor and acceptor RNAs. Primer extension to the end of the donor produces a 175-base full-length donor-directed DNA product. Strand transfers result in tension to the end of the donor RNA, allowing an additional 197-base full-length donor-directed products that includes the additional non-donor-directed DNA product. Strand transfers between regions of the HIV genome (Figs. 2). It is also possible for strand transfer to occur between two donor RNAs in the reaction. This type of transfer was not quantified in the experiments.

The percent transfer efficiency was defined as transfer DNA products (T)/transfer + full-length donor-directed products (F), times 100 ((T/(T+F)) × 100). The number reflects the proportion of DNA primers extended to the end of the acceptor versus those extended to the end of the donor. This representation of the data, as opposed to simply determining the gross level of transfer product, expresses transfer relative to total DNA extension. Therefore, differences in the total amount of primers extended with the various substrates used are compensated for.

Prediction of Secondary Structures of the RNA Substrates—RNAdraw (61) and Zuker’s RNA folding programs (62–64) were used to predict the secondary structures for the RNA substrates (both donor and acceptor) utilized in this report. Fig. 2 shows the structures of only the donor RNAs derived from the U3 LTR (A), gag-pol (B, referred to as GagPol substrate), RRE (C), pol-vif (2D), and env (E, referred to as Env substrate) regions. In each region, the acceptor RNAs were predicted to have structures similar to the donor (acceptor substrates not shown). A highly negative ΔG value and a high base pair melting temperature indicated the presence of strong structures. For example, the donor GagPol RNA had a predicted ΔG value of −45.3 kcal/mol, and the predicted stem loops persisted even at temperatures above 55 °C whereas the donor Env RNA showed a ΔG = −15.0, and the structure melted out completely above 55 °C. Although the RNA folding programs may not predict the structure of RNAs with 100% reliability, a reasonable estimate of the strength and characteristics of the RNA would be expected, especially because relatively small RNAs were used. In addition, the GagPol and RRE regions used have been shown previously (65–68) to possess strong stem loop structures similar to those predicted by the folding programs. For our purposes an exact rendering of the structures is not necessary; just a general idea of their relative strengths and the presence or absence of pause site is required. The latter can be evaluated from RT primer extension reactions.

Effect of NC Protein on Strand Transfer in Highly Structured RNA Substrates—In this section, autoradiograms of strand transfer assays conducted on three RNA substrates are shown in Fig. 3, and their corresponding graphical quantitations are shown in Fig. 4. The first pair of donor-acceptor RNA substrates we tested were from the U3 LTR region (corresponding to bases 9074–9265 of the HIV-1 genome as derived from plasmid pNL4–3 (56); see “Experimental Procedures”). This region folded to form a complex structure with several strong stem loops and a ΔG = −49.3 kcal/mol (Fig. 2A). A previous in vivo study identified a 24-base segment between bases 9158 and 9183 in this region as a potential hot spot for recombination (42). The study used two viral vectors based on different strains of HIV-1 to analyze recombination and obligatory strand transfers in HIV. The vector viruses were harvested...
from producer cells (CD4+ HIV-1 Env inducible cell line) and were used to infect target cells (CD4+ HeLa T4 cells). Of the 86-target cell clones analyzed, 11 clones were shown to undergo homologous recombination during minus strand U3 synthesis. DNA sequencing analysis of these clones revealed a 24-base region as a potential hot spot for recombination. The report also suggested that the region might have RNA secondary structures that induce pausing. Donor and acceptor RNA templates were generated from this zone and used in time course reactions. Fig. 3A is an autoradiogram of a PAGE showing the resolved DNA products. The bands at position 197, as indicated by the molecular ladder on the leftmost lane, are products of strand transfer from the donor to acceptor RNA. The bands at position 175 results from completed synthesis on the donor RNAs without any transfer events. Three distinct pause sites P1, P2, and P3 were observed between positions 65 and 90. P1 and P2 are within the 24-base region noted above, whereas P3 lies just outside at position 65. The major pause sites were mapped using DNA sequencing gels in which sequencing reactions on DNA were run using the primer from these reactions (data not shown). Interestingly, the characteristics of the position 65 site are strikingly similar to those of another pause site characterized previously (57, 69). Both sites are located within the stem of a strong stem loop structure. In both cases, RT stalled at a U residue located just behind a series of strong G-C pairs in the stem (Fig. 2B). Once again NC stimulated transfer from this region significantly (see Fig. 4A). In the absence of acceptor template the paused products persist even at 64 min, especially the 65-base products. Including NC protein without acceptor did not change the profile significantly. In the presence of acceptor template but absence of NC a decrease in the amount of the paused products was observed over time in comparison with reactions without acceptor. An increase in the amount of transfer product was also seen. The results are consistent with paused products transferring to the acceptor template and being extended. Paused products were "chased" more rapidly in the presence of NC and acceptor, and the level of transfer product also increased to a greater extent than in reactions with acceptor alone (Fig. 4A). Results here are in agreement with earlier reports (57, 70, 71) that have shown that paused DNAs can be focal points of strand transfer. The stimulatory effect of NC on recombination has also been supported by numerous experiments (15, 36, 46).

Transfer efficiency results from the above experiment are shown graphically in Fig. 4A. It is evident from this graph that NC enhances strand transfer significantly in this region of the genome. Each of the experiments was repeated two or more times to confirm the observed trends. Experiments presented are representative of the observations.

The second pair of donor-acceptor structured RNAs examined was from the gag-pol region (bases 2009–2200, ΔG = −45.30 kcal/mol) of the genome (Fig. 2B). This region includes the portion of genomic mRna where the programmed −1 ribosomal frameshifting event used to produce enzymatic proteins (RT, PR, IN) from the pol region occurs. The mechanism allows the virus to maintain a well regulated ratio of Gag proteins to GagPol proteins in infected cells, which is vital for efficient assembly of infectious particles. The segment of gag-pol region that we used included the heptanumeric "slippery site" (UUUUUUA sequence) and a downstream RNA secondary structure, both of which have been shown to promote efficient frameshifting (72, 73). Earlier literature has suggested that the RNA secondary structure is a simple stem loop (73, 74), but a recent report favors a more complex intramolecular triplex RNA structure (75). In either case, the structure seems to cause the ribosomes to pause and subsequently slip −1 base over the slippery site (Fig. 2B). Interestingly, this structure also caused the RT to pause as shown on Fig. 3B. This figure shows a time course reaction conducted on the GagPol substrates. The gel reveals a single strong pause site. Sequencing reactions mapped the site to an A residue at position 77 from the 3′ end of the donor RNA. In this case the A was also part of a strong stem structure and was followed by four G-C pairs. Once again NC stimulated transfer from this region significantly (see Fig. 3B and Fig. 4B), and production of transfer products was accompanied by an apparent chasing of the paused DNA product.

The third donor-acceptor pair of structured RNAs examined was from the RRE region (bases 7801–7992, ΔG = −59.40 kcal/mol) (Fig. 2C). The RRE has been characterized by Malim et al. (68) as a 234-base region to which the Rev protein can bind. Later on, Mann et al. (67) demonstrated that an extra 58 bases at the 5′ end and 59 bases at the 3′ end of the original 234 region (a total of 351 bases) was the complete biologically active RRE. The binding of Rev protein to this highly structured region is important for the nuclear export of unspliced and partially spliced HIV RNAs. A strong pause site was mapped to an A residue at position 72 in this template (see Fig. 2C and Fig. 3C).

As with the other structured RNAs examined above, NC enhanced strand transfer in the RRE region that was accompanied by chasing of the major paused DNA products. The predicted secondary structures seemed to obstruct RT as indicated by the presence of pause sites. Moreover, the disappearance of pause sites correlated with the presence of acceptor and was further enhanced by NC (Fig. 4C). The results suggest that the pause sites are probably focal points for strand transfers in these particular structured RNAs. Note that this does not

---

**Fig. 6.** Graph of efficiency of strand transfer versus time for the weakly structured RNA substrates. Graphs were derived from the experiments shown in Fig. 5. The substrates used are indicated above each graph. All the other markings are the same as in the legend for Fig. 4.
imply that all transfers from the RNAs occur from the pause sites, only that transfer is exaggerated from these positions.

Effect of NC Protein on Strand Transfer in Weakly Structured RNA Substrates—In this section, autoradiograms of strand transfer assays conducted on two RNA substrates are shown in Fig. 5, and their corresponding graphical quantitations are shown in Fig. 6. The first donor-acceptor pair of RNA substrates tested was from the \textit{pol-vif} region of the viral genome (bases 3419–3610, ΔG = −19.3 kcal/mol) (Fig. 2D). The experiment shown in Fig. 5A demonstrates that “strong” pause sites are absent on this donor. Surprisingly, the efficiency of strand transfer on this substrate in the absence of NC was significantly higher than for the more structured substrates described above. In addition, though NC did hasten the onset of strand transfer, the overall enhancement was less as compared with the highly structured RNAs (Fig. 6A).

Similar results were obtained with the second low structure donor-acceptor pair of RNA substrates from the \textit{env} region (bases 7079–7270, ΔG = −15 kcal/mol) (Fig. 2E). The region was predicted to have the weakest structure among the other four substrates, as shown by the ΔG value. This region also showed an absence of strong pause sites (Fig. 5B). Some weak pause sites were observed, but they disappeared quickly even in the absence of NC and acceptor RNA. Once again, the level of strand transfer was relatively high in the absence of NC and NC moderately increased transfer but to a lesser extent than on structured substrates (Fig. 6B).

Effect of Varying Concentrations of NC on Three RNA Substrates—We conducted strand transfer assays with increasing concentrations of wild type NC protein on three of the sub-
Two different mutants of NC protein were tested. Substrates with more structure were tested in a concentration-dependent manner and to a greater extent on the Env level. These observations show that NC enhances transfer in a concentration-dependent manner. Levels of efficiency at 1 μM NC were comparable for PolVif and U3 3′LTR regions whereas the PolVif region peaked at a lower level. These observations show that NC enhances transfer in a concentration-dependent manner and to a greater extent on RNAs with more structure.

**Effects of Mutant NC Proteins on Env and GagPol RNA Substrates**—Two different mutants of NC protein were tested to determine their effect on transfer from structured (GagPol) and non-structured (Env) regions. These were finger mutants in which one of the two NC zinc finger was replaced with the other (12). HIV NC has two non-identical zinc fingers, a N- and a C-terminal finger, denoted 1 and 2, respectively. In mutant 1.1, finger 1 replaces the finger 2 giving this protein two copies of finger 1. In 2.2, finger 2 replaces finger 1 giving this protein two copies of finger 2. Recent work in our laboratory has suggested that the fingers possess biased functional activities with finger 1 containing more nucleic acid unwinding activity and finger 2 having more annealing activity. Clearly, strand transfer is a complex process that requires both activities for maximal stimulation (see “Discussion”). However, by using substrates with vastly different levels of structure we reasoned that it might be possible to uncover the properties of the mutants. Presumably, transfer on the highly structured GagPol substrate would require more unwinding activity to destabilize structures in the transferring DNA and acceptor template. Transfer on the Env substrates should be less dependent on unwinding activity and finger 2 having more annealing activity. It was interesting that the low structure substrates showed somewhat better transfer than the others. This was in contrast to strand transfer on the highly structured GagPol substrate (see Fig. 8B and Fig. 9B). In this case 2.2 was the least stimulatory whereas 1.1 and wild type were comparable. The results are consistent with 2.2 retaining annealing activity but having less unwinding activity than 1.1 or wild type. This supports a role for finger 1 in unwinding. The role of finger 2 in annealing is less clear from these experiments, because the mutant without finger 2 (1.1) was nearly as good as wild type on both substrates.

**DISCUSSION**

In this study, an *in vitro* assay that simulates internal strand transfer events occurring during minus strand DNA synthesis was used. The substrates tested transfer over 150-base homologous regions and were chosen to span a wide range with respect to structural strength. This allowed a comparison of transfer on highly structured versus relatively non-structured RNAs in the presence or absence of NC. The level of transfer in the absence of NC was higher on substrates with low structure and without prominent DNA synthesis pause sites (see Figs. 5 and 6). NC only modestly stimulated transfer on these substrates. In contrast, highly structured substrates with prominent pause sites transferred relatively poorly in the absence of NC but were highly stimulated when NC was present (see Figs. 3 and 4). Much of the stimulation on the latter substrates appeared to result from the “chasing” of paused products into transfer products. Overall the results showed that strand transfer could occur efficiently by mechanisms that do not involve pausing, although when present, pause sites serve as focal points for transfer events. In addition, NC stimulated transfer of highly structured regions, probably by aiding in the unwinding of secondary structure, as mutant NC proteins with lower unwinding activity were less stimulatory (see Figs. 8 and 9).

It was interesting that the low structure substrates showed high levels of strand transfer even though very little pausing was observed. Both pause-driven transfer (57, 69–71, 76) and transfer driven by non-pause-related mechanism (based on interactions between template structures for example) have been proposed (77–81). In pause-driven strand transfer, paused DNA products are the substrates that ultimately are used to produce transfer products. The pause-driven mechanism proposes that when secondary structures stall the progression of RT, it creates an opportunity for increased degradation of donor RNA by RT RNase H activity (57, 82). The nascent DNA may be released from the donor, or the acceptor...
RNA may invade and displace the donor RNA. Eventually, the nascent DNA associates with the acceptor RNA, where synthesis continues. Although NC has been shown to reduce pausing, it may augment strand transfers under such a mechanism by accelerating interactions between the acceptor and complementary DNA. Recently, it has been shown that pausing not only results in transfer of the paused DNA product but can also promote transfers downstream of the pause site (70). In this case pausing apparently allows time for the acceptor to interact with the nascent DNA to form a trimeric complex consisting of the acceptor, donor, and DNA. The acceptor then “zips” up the complementary DNA and displaces the donor as RT continues extension of the paused DNA. This leads to transfer at a point beyond the pause site. Alternatively, other reports (79, 80) have proposed that interactions between the transcribing DNA and specific structures on the acceptor can trigger strand transfers in the absence of pausing. In this mechanism specific structures on the acceptor interact with the DNA after the formation of a trimeric structure. These interactions lead to displacement of the donor as the DNA binds to the acceptor. Ultimately, there may be several mechanisms by which strand transfer can occur. These could depend on local structures or sequences. The relative quantitative importance of the various mechanisms remains to be determined.

For those regions with high structure in our experiments, at least part of the transfer appeared to result from a standard pause-driven mechanism where DNAs stalled at the pause site transferred to the acceptor (see Fig. 3 and “Results”). The other regions (PolVif and Env) showed very little pausing, so some other type of mechanism is likely. One possibility is a non-pause-induced mechanism similar to the “zipping” mechanism proposed above. In the above mechanism pausing provides time for the formation of a trimer, which then resolves at a point downstream of the pause site. In low structure regions, pausing may not be necessary for trimer formation, because the lack of secondary structures that impede hybridization allows rapid association of the acceptor and nascent DNA. With this in mind, strand transfer could be viewed as a compromise between opposing processes. Pausing favors transfer by providing substrates that can readily transfer to the acceptor. However, the high degree of structure that is generally associated with pausing is not conducive to association between the acceptor and transferring DNA, thus opposing transfer. If most recombination events occurred by non-pause-induced zipping mechanisms, then recombination would likely be fairly homogeneously spread over the whole HIV genome and may even focus in regions with low structure. A recent report indicated that on average transfers occur—three times during production of a single provirus and that apparent “hotspots” probably exist (43). Whether there are local structures or pause sites in the hotspot regions will require a more detailed analysis of these regions.

The effect of two NC finger mutants on strand transfer was also examined. Several reports (17, 24, 83, 84) using NC finger deletion mutants to try and understand the role of specific protein domains have been published. Results show that the highly basic backbone of the molecule is responsible for a portion of the annealing activity (25), whereas zinc coordination by the finger domains is required for full unwinding activity (83, 84). As was noted earlier, recent results in our laboratory suggest that the two fingers may have disproportionate roles in annealing and unwinding with finger 2 possessing more of the former activity and finger 1 possessing more of the latter activity. Results reported here showed that mutant 1.1 (two copies of finger 1) enhanced strand transfer to approximately the same level as wild type on both the highly structured GagPol substrate and low structured Env substrate (Fig. 9, A and B). In contrast 2.2 (two copies of finger 2) showed greater enhancement than wild type on the Env substrate whereas stimulation was reduced on GagPol. These results are consistent with a role for finger 1 in unwinding as the loss of this finger resulted in reduced transfer in the more structured substrate. The results are at least partly consistent with a previous report (83) using these mutants. In that study the mutants were evaluated using a system to test strand transfer of the transactivation response containing minus strand strong-stop DNA (−ssDNA). In those experiments mutant 1.1 retained 50% of wild type activity, and 2.2 showed a complete loss of activity. The differences observed between these experiments and ours could have resulted from the substrates used. The −ssDNA substrate forms a very strong complex secondary structure, which may require a high degree of unwinding and annealing activity to overcome the strong structure. Also, −ssDNA has a tendency to self-prime, an effect that prevents transfer but is inhibited by NC (47, 85). Therefore this substrate may be very sensitive to small changes in NC that have an effect on self-priming and/or annealing and unwinding. Overall, however, both sets of experiments support a role for finger 1 in helix destabilization.

Results presented here indicate that the role of NC in recombination is likely dependent on the structure of the region from where strand transfer is occurring. Unwinding activity is especially important in this process, because regions with high structure appear to transfer relatively poorly unless NC is present. Consistent with recent findings (81), results also showed that strand transfer could occur in the absence of strong polymerase pausing, perhaps because of more rapid association of the acceptor and DNA in non-structured regions. The emerging story for retrovirus recombination is one in which events can occur by several potential mechanisms. Further experiments will be required to uncover the details of the various mechanisms.

Acknowledgments—We thank Genetics Institutes for the kind gift of HIV-RT and the AIDS Research and Reference Reagent Program for plasmid pNL4–3. We also thank Dr. Robert Gorelick from Science Applications International Corp., Frederick, MD for NC mutant proteins and Dr. Charles McHenry from the University of Colorado for the plasmid clone for wild type NC overexpression.

REFERENCES
1. Karpe, R. L., Henderson, L. E., and Oroszlan, S. (1987) J. Biol. Chem. 262, 4961–4967
2. Frankel, A. D., and Young, J. A. (1998) Annu. Rev. Biochem. 67, 1–25
3. Titech, R. K., Cheng, Y. Y., Yin, F. F., and Erickson-Viitanen, S. (1991) J. Virol. 65, 922–930
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, 1856–1965
5. Mely, Y., De Roquigny, H., Morellet, N., Roques, B. P., and Gerard, D. (1996) Biochemistry 35, 5175–5182
6. Henderson, L. E., Copeland, T. D., Sowder, R. C., Smythers, G. W., and Oroszlan, S. (1981) J. Biol. Chem. 256, 840–8406
7. Bess, J. W., Jr., Powell, P. J., Issaq, H. J., Schumack, L. J., Grimes, M. K., Henderson, L. E., and Arthur, L. O. (1992) J. Virol. 66, 840–8407
8. Mely, Y., Julian, N., Morellet, N., De Roquigny, H., Deng, C. Z., Piemont, E., Roques, B. P., and Gerard, D. (1994) Biochemistry 33, 12085–12091
9. Morellet, N., De Roquigny, H., Mely, Y., Julian, N., Demene, H., Ottmann, M., Gerard, D., Darlix, J. L., Fournier-Zaluski, M. C., and Roques, B. P. (1994) J. Mol. Biol. 235, 287–301
10. Morellet, N., Julian, N., De Roquigny, H., Maiquet, B., Darlix, J. L., and Roques, B. P. (1992) EMBO J. 11, 3959–3965
11. South, T. L., Blake, P. B., Hare, D. R., and Summers, M. F. (1991) Biochemistry 30, 6342–6349
12. Gorelick, R. J., Chabot, D. J., Rein, A., Henderson, L. E., and Arthur, L. O. (1993) J. Virol. 67, 4027–4036
13. Horlick, R. J., Gagliardi, T. D., Bosche, W. J., Wiltrout, T. A., Coren, L. V., Morellet, N., De Rocquigny, H., Jullian, N., Demene, H., Ottmann, M., Gerard, D., Darlix, J. L., Fournier-Zaluski, M. C., and Roques, B. P. (1994) J. Biol. Chem. 269, 92–104
14. Darlix, J. L., Lapadat-Tapolsky, M., de Roquigny, H., and Roques, B. P. (1995) J. Mol. Biol. 254, 523–535
15. De Stefano, J. J. (1996) J. Biol. Chem. 271, 16350–16356
16. Gue, J., Wu, T., Anderson, J., Kane, B. F., Johnson, D. G., Gorelick, R. J., Henderson, L. E., and Levin, J. G. (2000) J. Virol. 74, 8980–8988
Evidence for the Differential Effects of Nucleocapsid Protein on Strand Transfer in Various Regions of the HIV Genome
Suchitra S. Derebail, Megan J. Heath and Jeffrey J. DeStefano

J. Biol. Chem. 2003, 278:15702-15712.
doi: 10.1074/jbc.M211701200 originally published online February 20, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M211701200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 83 references, 54 of which can be accessed free at http://www.jbc.org/content/278/18/15702.full.html#ref-list-1