Published in final edited form as:

J Mol Biol. 2009 April 24; 388(1): 30–47. doi:10.1016/j.jmb.2009.02.029.

Mechanism Analysis Indicates That Recombination Events In HIV-1 Initiate And Complete Over Short Distances, Explaining Why Recombination Frequency is Similar in Different Sections of the Genome

Sean T. Rigby¹, April E. Rose¹, Mark N. Hanson¹, and Robert A. Bambara^{1,2,*}

¹Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642, USA

²Cancer Center, University of Rochester Medical Center, Rochester, NY 14642, USA

Abstract

Strand transfer drives recombination between the co-packaged genomes of HIV-1, a process that allows rapid viral evolution. The proposed invasion-mediated mechanism of strand transfer during HIV-1 reverse transcription has three steps: invasion of the initial or donor primer-template by the second or acceptor template, propagation of the primer-acceptor hybrid, and then primer terminus transfer. Invasion occurs at a site at which the RT RNase has created a nick or short gap in the donor template. We used biochemical reconstitution to determine the distance over which a single invasion site can promote transfer. The DNA-primed RNA donor template used had a single stranded precreated invasion site (PCIS). Results showed that the PCIS could influence transfer twenty or more nucleotides in the direction of synthesis. This influence was augmented by viral nucleocapsid protein (NC) and additional reverse transcriptase (RT) ribonuclease H (RNase H) cleavage. Strand exchange assays were performed specifically to assess the distance over which a hybrid interaction initiated at the PCIS could propagate to achieve transfer. Propagation by simple branch migration of strands was limited to 24 – 32 nucleotides. Additional RNase H cuts in the donor RNA allowed propagation to a maximum distance of 32 – 64 nucleotides. Overall, results indicate that a specific invasion site has a limited range of influence on strand transfer. Evidently, a series of invasion sites cannot collaborate over a long distance to promote transfer. This result explains why the frequency of recombination events does not increase with increasing distance from the start of synthesis, a characteristic that supports effective mixing of viral mutations.

Keywords

HIV-1; reverse transcriptase; strand transfer; recombination; hybrid propagation

© 2009 Elsevier Ltd. All rights reserved.

* Corresponding author .

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

Recombination during reverse transcription is a source of genetic diversity in HIV-1. During reverse transcription mutations are frequently generated^{1; 2; 3; 4; 5; 6; 7; 8}. Differing co-packaged RNA genomes can recombine advantageous genetic code, resulting in improved viral fitness^{9; 10; 11; 12; 13; 14; 15}. Reverse transcriptase (RT) is the enzymatic protagonist of reverse transcription. The two catalytic activities of RT are polymerization and RNase H, i.e. cleavage of RNA strands that are annealed to DNA^{16; 17; 18; 19}. Through these two activities RT reverse transcribes dsDNA from the genomic ssRNA templates. Successful completion of reverse transcription requires strand transfer, a process by which the nascent DNA switches between RNA templates²⁰. Through strand transfer sequence variations between copackaged RNA genomes can be combined in the dsDNA product.

Strand transfer recombination can occur at any time during minus strand synthesis, as described by the “dynamic copy choice” model, and recombination is dependent on the ratio between rates of RNA cleavage and DNA synthesis^{21; 22; 23}. A ratio in favor of RNA cleavage produces higher efficiency of strand transfer, while less cleavage and an increased rate of polymerization favors a lower transfer efficiency. Results of many experiments performed both *in vivo* and *in vitro* indicate that cleavage of the RNA leads to an “invasion-mediated” mechanism of strand transfer; a process in which a second RNA invades gaps in the initial RNA template to hybridize with ssDNA^{24; 25; 26; 27; 28}.

During minus strand synthesis, the RT degrades the RNA concomitantly with polymerization, though cleavages are not made as frequently as nucleotides are incorporated, leaving fragments of RNA still annealed to the nascent DNA^{29; 30; 31; 32}. Additional RT molecules further degrade the remaining fragments independently of polymerization^{30; 31; 33; 34; 35; 36; 37; 38; 39}. RNA cleavage leads to the emergence of invasion sites, sections of ssDNA that are exposed for the second, intact RNA to invade and hybridize^{27; 28}. Results obtained *in vitro* suggest that creation of the invasion site is the rate-limiting step of the invasion-mediated mechanism of strand transfer⁴⁰. After invasion, the RNA-DNA hybrid must propagate to the 3'-terminus of the nascent DNA to drive a terminus transfer, which completes the template switch. Hybrid propagation can proceed by the successive basepairing of nascent DNA with the second RNA, known as branch migration. Heath et al. demonstrated transfer-related branch migration in strand exchange reactions in the absence of RT⁴¹.

Branch migration is one of two possible mechanisms of hybrid propagation. The other is a proximity-based mechanism. The proximity effect is achieved through an invasion upstream of the DNA 3'-terminus, which increases the local concentration of the second RNA template around the nascent DNA. The proximity effect can engender terminus transfer as suggested by the work of Song et al.⁴².

In a way similar to the proximity mechanism, nucleocapsid protein (NC) also stimulates annealing by increasing local concentrations of complementary nucleic acid by way of aggregation^{43; 44; 45}. NC, a 55 amino acid product of gag gene expression and processing, has two zinc fingers that allow it to interact with and chaperone nucleic acids, with a preference for ssRNA^{43; 46; 47; 48}. Not only does NC promote aggregation, but it also destabilizes weak secondary structures^{49; 50; 51; 52; 53}. Accordingly, NC increases strand transfer efficiency, in part, through facilitating the invasion step^{47; 48; 54; 55; 56; 57; 58}. Additionally, NC promotes strand exchange so long as the product shares a greater region of complementarity than the reactant^{41; 51; 52}.

It has been reported that 3 – 30 strand transfer events occur per genome per replication cycle^{59; 60; 61; 62; 63}. During reverse transcription, as an RT extends the primer terminus farther away from the primer binding site (PBS) the polymerizing RT and assisting RTs cleave the

RNA template creating an ever increasing number of potential invasion sites, i.e. the number of potential invasion sites increases proportionally with the length of cDNA extension. This would predict that transfer efficiency at any position of primer extension would be proportional to the distance of synthesis from the primer to that position. However, Jetzt et al. reported that recombination efficiency across the length of the HIV-1 genome varied between regions, but did not increase with distance from the PBS⁶⁰. This result suggests that some factor limits efficiency of strand transfer despite increasing opportunities for invasion. The study by Jetzt et al. is consistent with the expectation that the farther the primer terminus is removed from an invasion site, the less likely the RNA-cDNA hybrid created by invasion can propagate quickly and efficiently enough to catch the DNA terminus and complete the transfer⁶⁰. That is, efficiency by which hybrid propagation can contribute to transfer is expected to decrease substantially with distance. Considering the likelihood of limitations on the distances between invasion and terminus transfer, and the limited number of transfer events, it is probable that transfer mechanisms are determined by local structures and conditions. The factors that determine hybrid propagation rate and efficiency have heretofore been largely unexplored.

Measurements of strand transfer made *in vitro* suggest that distances from invasion sites to points of terminus transfer can be separated by dozens of nucleotides^{28; 40; 64; 65; 66}. However, each substrate system employed has had several potential sites of invasion excluding precise measurement of the distance over which a specific invasion site can promote terminus transfer. Studies have investigated the influences of NC on strand exchange reactions with a specific invasion site over shorter lengths, but the limitations of propagation distance with respect to specific invasion sites are undefined^{41; 51; 52}. Herein we use specially designed RNA and DNA transfer substrates *in vitro* to determine how individual invasion sites contribute to the invasion-mediated mechanism of strand transfer, with specific focus on the limits and influences on the apparently crucial intermediate step of hybrid propagation. In particular, we investigate the roles of the RNase H activities of RT and the strand exchange properties of NC.

Results

Our objectives were threefold; to design and use a reconstituted system able to test directly for the influence of a specific invasion site in strand transfer; to determine whether the hybrid propagation step of transfer is limited by distance; and to establish whether hybrid propagation is affected by the strand exchange properties of NC and the RNase H activity of RT.

Pre-created Invasion Site (PCIS) Substrates Used in Hybrid Propagation Experiments

PCIS substrates have a single stranded region that allows rapid, efficient acceptor invasion at a specific location⁴⁰. The PCIS substrates used in this study were comprised of a DNA primer annealed to either a DNA or RNA donor template in which the primer had a 5'-overhang of 20 nt (Figure 1). This region of the ssDNA-primer is the PCIS, and has a G/C content of 45 % to promote invasion. Substrates are labeled either – PCIS or +PCIS indicating the absence or presence of the 20 nt 5'-overhang, respectively. An acceptor template, A_m, was designed to have full complementarity with the primer including the region of the PCIS with the exception of 3 nt that created a mismatch between the primer 3'-terminus and the acceptor template (Table 1). The mismatch was introduced to prevent polymerization on the acceptor should the primer exchange from the donor to the acceptor before initiation of synthesis on the donor. We observed that the 3 nt mismatch inhibited template directed synthesis in reactions with primed acceptor, and no donor (data not shown). A U33A base substitution was also introduced for the strand exchange reactions described elsewhere (see *Materials and Methods*). The +PCIS substrate is designed to resemble the strand transfer reaction intermediate just after creation of the invasion site.

The PCIS substrates offer unique advantages for characterizing the transfer mechanism. Comparing transfer efficiency from substrates having *versus* lacking the PCIS allows a confident determination of the role of invasion at that specific site on subsequent transfer. Since the substrates are designed so that the primer terminus must progress well downstream of the PCIS before transfer can occur, evidence that transfers initiate at the PCIS and complete later would provide the strongest model *in vitro* to date supporting the proposed invasion-mediated transfer mechanism.

Use of these substrates also facilitates assessment of the effects of two reaction components that we predicted to influence hybrid propagation and terminus transfer. The first is NC, which has been reported to promote strand exchange. The second is RT RNase H, which is anticipated to allow the initial DNA-RNA hybrid to expand by cutting the donor RNA so as to destabilize its interaction with the cDNA. The effects of RNase H can be eliminated experimentally through the use of the E478Q mutant RT, which lacks RNase H function.

Effects of the PCIS and NC on Transfer

We first questioned whether the presence of the PCIS influenced transfer. We used the primers -PCIS₂₀ or +PCIS₂₀, which were complementary with the last 20 nt of the 3'-end of the donor template, and had 3 nt mismatch with the acceptor RNA, A_m, at nucleotide positions 18 – 20 (Table 1). Results indicated that the PCIS generally increased transfer efficiency (see *Materials and Methods* for determination of transfer efficiency) whether or not NC or RNase H were present (Fig. 2). The transfer efficiency of reactions with E478Q RT and no PCIS was a very low 0.6 %, but was significantly increased with the PCIS by as much as 6-fold to 3.6 % (Fig. 2b,c and Table 2). Comparatively, the increase of transfer efficiency when NC and PCIS were combined in the presence of E478Q was approximately 19-fold to 11.4 %.

We next asked whether NC could substitute for the PCIS. In a reaction with E478Q RT on a substrate without the PCIS, addition of NC increased the transfer efficiency a mere 1.5-fold to 0.9 %. Evidently the PCIS had a more influential role in promoting strand transfer than NC when RT-RNase H activity was absent.

To further clarify the effects of a PCIS and NC on strand transfer in the absence of RNase H the previous reactions were repeated with WT RT and a DNA donor template instead of E478Q and an RNA donor template. The transfer efficiency for these reactions without a PCIS and NC was 0.4 % (Fig. 2b,c). Much like the E478Q reactions, the low transfer efficiency of the DNA donor reactions was increased with the addition of a PCIS by 3.8-fold to 1.5 % (Table 2). Again, NC had little effect on strand transfer without the PCIS, altering the transfer efficiency only to 0.3 %, but stimulated transfer 7.5-fold to 3.0 % when the PCIS was present. While the trend of the results with WT RT and a DNA donor is similar to E478Q RT and an RNA donor, the transfer efficiencies were generally lower, presumably because of the characteristics of the DNA-DNA primer-donor helix. Results with both reaction systems support the conclusion that the PCIS is critical for efficient transfer for a template system in which RNase H does not cut the donor template. Moreover, both systems indicate that NC functions primarily through facilitating the role of the PCIS in transfer. Very likely, this includes promoting acceptor-cDNA primer interaction at the PCIS and subsequent strand exchange to complete transfer. This assumption is in agreement with evidence showing that substrates with a pre-created invasion site have an increased rate of strand transfer with increasing acceptor concentration, and that NC enhances the formation of stable acceptor-primer hybrids at the pre-created invasion site ^{40; 41}.

Effects of RNase H on Transfer

To examine the influence of RNase H on strand transfer with the PCIS substrates, the above experiments were repeated using WT RT with the RNA donor substrates. Even without the PCIS and NC the strand transfer reactions yielded a distinctly higher 5.1 % efficiency with RNase H (Fig. 2b,c). With the introduction of the invasion site, transfer increased 2-fold to 10.5 %. Without the PCIS, NC increased transfer efficiency 1.7-fold to 8.6 %. In combination, the PCIS and NC stimulated transfer efficiency 3-fold to 15.6 %. Clearly, the introduction of RNase H activity markedly reduced the percentage impact of both the stimulation by the PCIS, and combined PCIS and NC stimulation of strand transfer. The percent effect of NC alone, however, remained similar to results without RNase H. A reasonable conclusion is that cuts made by the RT RNase H partially substitute for the PCIS and NC. The RNase H cuts are likely to create alternative invasion sites, and also weaken the donor-primer interaction to reduce the energy barrier to strand exchange from the PCIS and every invasion site, similar to the effects of NC.

A second approach to examining the effects of RNase H activity on invasion-mediated strand transfer was to repeat the previously described E478Q or WT RT reactions, but with exogenous RNase H activity in the form of *E. coli* RNase H. The addition of exogenous RNase H (0.08 U/ μ l) increased transfer efficiency with substrates lacking the PCIS from 0.6 % to 1.8 % (Fig. 2d,e). In reactions with the PCIS, exogenous RNase H increased transfer efficiency from 5.3 % to 6.8 % and 13.8 % for 0.02 and 0.08 U/ μ l, respectively. Upon addition of NC, transfer efficiency without the PCIS was increased from 0.9 % to a significant 6.0% and 3.6% for 0.02 and 0.08 U/ μ l, respectively. Addition of NC to reactions with the PCIS slightly altered the transfer efficiency from 12.9 % to 13.1 % and 11.5 % for 0.02 and 0.08 U/ μ l, respectively. The –PCIS to +PCIS ratio for E478Q reactions lacking exogenous RNase H was 3.8 without NC, and 13.6 with NC (Fig. 2f). The addition of exogenous RNase H dropped the –PCIS to +PCIS transfer ratio to near the levels seen with WT RT. Inclusion of exogenous RNase H, as with use of WT RT, reduced the influence of the PCIS in strand transfer reactions. Interestingly, NC stimulated transfer efficiency in reactions lacking the PCIS only when RNase H was present, either exogenous or with WT RT, perhaps through enhancing interactions between the acceptor and primer at invasion sites created during the reactions by RNase H activity. We suspect that in WT RT reactions the PCIS was the more commonly used invasion site as compared to invasion sites created by RNase H activity because there was less stimulation by NC in exogenous RNase H-containing reactions with WT than with E478Q.

Most permutations of reaction conditions with E478Q RT and *E. coli* RNase H did not yield transfer efficiencies similar to WT levels. These results suggest that there is some advantage to transfer efficiency of combining polymerization and RNase H functions in the same molecule, or that the ratio of the rates of polymerase to RNase H activity is critical for efficient strand transfer.

Distance Effects on Hybrid Propagation

As discussed above, the fact that recombination in HIV-1 does not increase with synthesis distance from the PBS suggests that the invasion site can affect transfer over only a limited distance. To address this issue mechanistically, we tested the distance-related restrictions on hybrid propagation by measuring the efficiency of template strand exchange over time, as occurs during the transfer reaction, using a substrate design that simulates primer elongation to a specific distance from the PCIS. The approach was to construct substrates in which length of the primer hybridized to the donor RNA was altered from the original 20 nt to 17, 24, 32 and 64 nt (Fig. 3a, 4a, 5a, 6a). Each acceptor RNA to be used was constructed with a base substitution, compared to the donor RNA sequence, 1 nt after the 3' terminus of the tested primer. Time-dependent completion of strand exchange was assessed by the efficiency of

incorporation of the nucleotide specific for base pairing to the acceptor RNA. This was accomplished by having only the appropriate dNTP for incorporation after exchange present as a reactant. To assess the influence of the 3 nt mismatch in strand transfer reactions with the primers -PCIS₂₀ and +PCIS₂₀ in Figure 2, each propagation distance was measured with acceptor templates that either contained the 3 nt mismatch or were perfectly complementary with the primer.

For nucleotide incorporation to represent the strand exchange rate, incorporation must be fast relative to strand exchange. To measure the rate of incorporation, control reactions were performed with the primer pre-annealed to the acceptor RNA. A comparison of exchange rate to extension rate on the acceptor demonstrated that nucleotide incorporation was not determining the rate of exchange plus incorporation. Therefore, in the actual reaction, nucleotide incorporation was an appropriate quantitative indicator of strand exchange rate. Reactions were also performed with the primer pre-annealed to the donor template, and no acceptor template present, to control for non-template-directed dNTP incorporation. Exchange efficiencies rising above this background control are therefore a result of strand exchange and not misincorporation. The efficiencies of incorporation on the donor and acceptor templates are represented along with the exchange efficiency on a graphical axis of exchange/extension efficiency so that the three values can be directly compared (Fig. 3 – 6).

Reactions were performed using substrates either having or lacking the PCIS. This allowed us to quantitate the effect of the PCIS on the rate of exchange of the primer from the donor to the acceptor template. For example, some reactions with a PCIS and E478Q RT show clear evidence of strand exchange by branch migration, specifically deriving from invasion at the PCIS, not observed in same reactions lacking a PCIS. Each distance has revealed unique distance-related characteristics of the hybrid propagation step.

Characteristics of Each Propagation Distance

Reactions with primer set, -PCIS₆₄ and +PCIS₆₄, showed that at the longest propagation distance tested, 64 nt, transfer efficiency with E478Q RT remained at background, even in the presence of the PCIS and NC (Fig. 3c,e). Yet, with WT RT, although the PCIS alone had no distinguishable influence on transfer efficiency, addition of NC to the PCIS substrate increased the transfer rate and efficiency (Fig. 3b,d). These results show that RNase H activity was required for efficient transfer when a distance of 64 nt separated the invasion site and primer terminus. Moreover, NC influenced strand transfer in PCIS reactions only when RNase H activity was present. These observations imply that once the hybrid propagation length reaches approximately 64 nt, hybrid propagation over the full distance rarely occurs, and that creation of new invasion sites closer to the primer 3'-terminus is necessary for efficient strand transfer. The differences between reactions with and without the primer-acceptor 3 nt mismatch are minimal, which suggests that the mismatch had little influence on strand exchange over 64 nt.

Reactions with -PCIS₃₂ and +PCIS₃₂ revealed the distance, approximately 32 nt, at which the PCIS moderately influenced strand transfer efficiency, but only with WT RT. Just as with the 64 nt propagation distance, strand exchange at the 32 nt distance with the E487Q RT remained at background, even with the PCIS and NC for reactions with the 3 nt mismatch (Fig. 4c,e). Complementary reactions without the mismatch displayed a moderate rise of strand exchange above background with PICS, NC and E478Q. This difference is likely due to altered acceptor structure, although it is possible that there is inhibition of branch migration beyond the 3 nt mismatch. Our purpose for comparing strand exchange efficiency with and without the primer-acceptor mismatch was to determine whether the mismatch inhibited hybrid propagation in the strand transfer reactions. It appears that a mismatched region of three nucleotides has little effect on strand exchange across a distance of 32 nt, and probably did not significantly influence the strand transfer reactions.

Similar to the 64 nt reactions, NC increased exchange efficiency with WT RT, however, unlike the longer distance, at 32 nt the +PCIS reactions had a higher transfer efficiency than –PCIS reactions (Fig. 4b,d). Exchange efficiencies with the 3 nt mismatch-acceptor templates, WT RT, but no NC, were 39 % and 20 % for +PCIS₃₂ and –PCIS₃₂, respectively. For the similar reactions with complementary primer and acceptor, exchange efficiencies were 22 % and 14 % for +PCIS₃₂ and –PCIS₃₂, respectively. Strikingly, exchange efficiencies with NC achieved the high values of 70 % and 45 % for +PCIS₃₂ and –PCIS₃₂, respectively. For the similar reactions with complementary primer and acceptor, exchange efficiencies were 48 % and 29 % for +PCIS₃₂ and –PCIS₃₂, respectively. Clearly, the PCIS was more influential in strand exchange with the 32 nt distance than with a 64 nt hybrid propagation distance. Results with E478Q RT revealed minimal influence of the PCIS on exchange efficiency with the aid of NC. Therefore, RNase H activity still greatly facilitated efficient exchange with a 32 nt propagation distance, but was not absolutely required. At 32 nt branch migration alone, though highly inefficient, could be a pathway in the mechanism of strand transfer.

Results from strand exchange experiments with primers –PCIS₂₄ and +PCIS₂₄ reveal that branch migration did not occur over a distance of 24 nt without RNase H activity and the PCIS (Fig. 5). NC could stimulate efficient branch migration without RNase H activity so long as the PCIS was present. Upon introduction of RNase H activity with the WT RT, exchange efficiency rose above background even without the PCIS, although the PCIS significantly increased transfer efficiency in the presence of NC. A comparison of the results from the 32 nt strand exchange experiments to the 24 nt distance experiments implies that with decreasing distances of hybrid propagation NC increasingly promotes invasion-mediated branch migration.

It was not possible to assess strand exchange over 24 nt with mismatched substrates because nucleotide incorporation from the primer on the acceptor template was approximately an order of magnitude slower than in all other exchange reactions (data not shown). This was likely because the 3 nt mismatch at positions 18 – 20 destabilized primer-acceptor annealing at positions 21 – 24, which inhibited rapid extension at position 25.

The primers –PCIS₁₇, +PCIS₁₇, –PCIS₁₈, and +PCIS₁₈ are all complementary with the last 17 nt of the D_{RNA} 3'-end (Fig. 1; Table 1). The primer pair, –PCIS₁₇ and +PCIS₁₇, are complementary up to position 17 on A₁₇, and the pair –PCIS₁₈ and +PCIS₁₈ are complementary up to position 18 on A_m (Fig. 1, 6a; Table 1). These are 17 nt exchange reactions because the length of donor that must be displaced by the hybrid propagation reaction is the same for both substrates. The 17 nt exchange results revealed that strand transfer can occur by strand exchange alone even with E478Q. Even with this substrate, however, RNase H increases transfer efficiency. E478Q reactions had an exchange efficiency significantly above background only when the PCIS was present. NC increased the transfer efficiency of only +PCIS reactions with the 3-nt mismatched acceptor, Am2, and both +PCIS and –PCIS reactions with the complementary acceptor, Am9 (Fig. 6c,e). With WT RT, PCIS-mediated reactions were more efficient than –PCIS reactions, and NC increased the transfer efficiency of both –PCIS and +PCIS reactions (Fig. 6b,d). Differences between reactions with the acceptors Am2 and Am9 are likely from differences in template structures since the primers are entirely complementary to their respective acceptors. Results suggest that when the hybrid needs to propagate only 17 nt, the approximate distance separating the RNase H and polymerase domains of RT, strand exchange is facile, even in the absence of RNase H, but always requires the PCIS.

Discussion

The proposed invasion mechanism for strand transfer in retroviruses such as HIV-1 involves the steps of acceptor invasion, RNA-DNA hybrid propagation, and ultimate terminus transfer of the DNA from the donor to the acceptor RNA template. Depending on the relative rates of cDNA synthesis and hybrid propagation, positions of invasion and terminus transfer could be separated by considerable distances. Primer extension from the PBS during reverse transcription *in vivo* is expected to create progressively more potential invasion sites for transfer. However, transfer events previously measured across the HIV-1 genome showed no significant increase in transfer efficiency the farther the DNA 3'-terminus was extended from the PBS⁶⁰. These results imply a distance-related limitation on efficient invasion-mediated strand transfer. Here we report the development and use of a model system *in vitro* designed to measure the distance over which a specific invasion site can promote strand transfer, and the effects of substrate structure and protein components of the reaction. The strand exchange step of transfer was previously shown to require a single stranded region in the primer-donor substrate to act as a site for invasion by the acceptor^{40; 41; 52; 66; 67; 68; 69}. Heath *et al.* examined the effects of NC, invasion site size, and primer-acceptor complementarity on the strand exchange reaction. It was observed that longer invasion sites generally lead to higher transfer efficiency and that strand exchange could propagate through 25 nt with as little as 10 nt for invasion⁴¹. Also, NC was demonstrated to inhibit strand exchange for acceptor-primer substrates with mismatches in the invasion region, and promote the exchange reaction by stimulating the interaction of complementary acceptor and primer sequences⁴¹. We have expanded on these results to present a comprehensive biochemical analysis of the strand transfer reaction. It provides the best evidence so far that strand transfers actually occur by a multi-step invasion-mediated mechanism. The force of evidence comes from the observation that the presence of a specific pre-created invasion site (PCIS) increased the efficiency of a primer terminus transfer occurring well downstream with respect to synthesis. Moreover, the analysis defined distance limitations on invasion-initiated hybrid propagation leading to transfer, and revealed the significance of combinations of a PCIS, NC and RT RNase H on efficiency of transfer.

A comparison of results of RT-mediated transfer assays using substrates having or lacking a PCIS, and without NC or RNase H, clearly demonstrated the need for an invasion site to allow transfer. The experiments revealed that, whether or not NC is present, it is essential to have an invasion site for efficient primer-acceptor branch migration through a stable primer-donor hybrid. This suggests that the acceptor template simply cannot invade a fully double stranded primer-donor substrate. Substrates with the PCIS allow some transfer, which is greatly enhanced by the presence of NC. In view of the well-known strand exchange properties of NC^{43; 70}, it is logical to conclude that NC promotes both a stable initial interaction of the acceptor with the PCIS, and an exchange of the donor and acceptor strands on the cDNA primer. This conclusion is fully consistent with the earlier observations of NC stimulation of strand exchange^{41; 43; 51}.

Reactions using WT RT instead of the RNase H-deficient E478Q mutant RT, add an additional level of complexity. RNase H-containing RT, in the presence of NC, considerably increases transfer efficiency on substrates both having and lacking the PCIS. A reasonable interpretation is that the RNase H creates additional sites where the acceptor can invade. These would be generated throughout the donor RNA-cDNA hybrid region irrespective of the initial presence of a PCIS. Indeed, previous studies have shown that RNase H activity does promote initiation of strand transfer by creating invasion sites^{2; 25; 26; 28; 40; 66; 67; 68; 69; 71}.

We envision that the ability of the RNase H to make cuts in the hybrid region could result in additional phenomena. One possibility is that cuts in the donor template in the hybrid region,

would destabilize cDNA-donor helical stability. This would allow a cDNA-acceptor interaction at the PCIS to propagate more rapidly into the cDNA-donor region. However, if this were the primary effect of RNase H, we would expect that the relative amount of transfer observed would increase for substrates having *versus* lacking a PCIS. In fact the relative effectiveness of the PCIS is lessened by the presence of RNase H.

Moreover, the presence of both the PCIS and RNase H might promote transfer by an “acceptor proximity mechanism”. In this process, the acceptor would initially interact with the PCIS. This interaction would raise the local acceptor concentration at downstream sites of RNase H cleavage. The proximity of the acceptor to these sites would promote additional interactions that would result in facile exchange of the acceptor for the donor. However, this mechanism would predict a greater relative effectiveness of substrates with the PCIS, which was not seen. While the destabilization and proximity effects are likely to make some contribution to transfer, the major effect of the RNase H appears to be the creation of additional invasion sites that can be used in addition to the PCIS.

Even though RNase H reduced the impact of the PCIS on strand transfer, both NC and the PCIS still increased transfer efficiency in reactions with WT RT, and they were more effective at stimulating transfer in combination than separately. The stimulation by the PCIS and NC indicate that the branch migration or proximity mechanisms are still occurring in the presence of WT RT. Therefore, in the presence of the PCIS, NC and RNase H several transfer mechanisms collaborate to produce maximum transfer product.

If the primary effect of the presence of RNase H is to create usable invasion sites, the fact that RNase H greatly reduces the impact of the PCIS indicates that invasion sites closer to the DNA primer terminus are much more effective than farther invasion sites. It would follow that the effectiveness of the PCIS at promoting transfers falls off precipitously with the distance between the PCIS and the DNA primer terminus. This means that most transfers that occur by an invasion mechanism, employing branch migration or acceptor proximity propagate over a relatively short distance. While NC appears to facilitate branch migration, lengthening the effective range of an invasion site, the ultimate range must be confined. The effect of continuous RNase H activity, would be to create a series of invasion sites from which transfers can initiate. This would produce hotspots of recombination across the HIV-1 genome, but no increase of recombination with increasing distance from the PBS. In fact this is exactly what has been observed *in vivo*⁶⁰.

To better define the distance limitations on hybrid propagation in strand transfer we performed strand exchange experiments with substrates in which the length of the primer-donor hybrid was progressively altered. Each of the lengths, 17, 24, 32, and 64 nt, revealed valuable insights into the hybrid propagation mechanism. Exchange over the short distance of 17 nt occurred efficiently without RNase H, but was still accelerated by the presence of RNA cleavage. Without RNase H, the PCIS was essential for exchange. Even with RNase H, the PCIS had a positive influence. As discussed above, the ability of RNA cleavage to modestly accelerate branch migration might have been masked in the transfer experiments by the more dominant phenomenon of invasion site creation. Overall, results from reactions with the 17 nt duplex revealed that continuous migration of uncut branches is likely the dominant means of connecting an initial invasion with a successful primer terminus transfer (Fig. 7).

Branch migration through 24 nt of uncut strands was only efficient with the assistance of NC, and when the 3'-terminus and the PCIS were separated by 32 nt, migration of uncut strands was practically ineffective for strand exchange. This conclusion is based on the requirement for NC to rescue transfer in the presence of E478Q RT at 24 nt, and the limited ability of NC to rescue transfer at 32 nt. In reactions with WT RT, the PCIS and NC both had a positive effect

on the exchange reaction. This outcome suggests two possible mechanisms, discussed above, by which the contribution of the PCIS to strand exchange requires RNase H cleavage. Either the RNase H weakens donor RNA-cDNA interactions so that there can be branch migration of the initial PCIS-acceptor hybrid, or the RNase H cleavage sites provide secondary points of acceptor interaction by proximity, after the initial PCIS-acceptor hybrid formation. This conclusion is consistent with our results showing that the presence of RNase H significantly improves transfer efficiency on a substrate with a PCIS.

As expected, the combination of E478Q RT and NC remained ineffective at promoting strand exchange when the PCIS and primer terminus were separated by 64 nt. However, exchange remained possible with WT RT and was promoted by NC. The notable feature of exchange reactions with this substrate is that there is no significant influence of the PCIS. Evidently, exchange proceeds completely from invasions initiated at RNase H cleavage sites beyond the PCIS (Fig. 7). Therefore, the 64 nt separation represents a distance completely beyond the range of hybrid propagation for strand transfer by any mechanism.

Based on our observations, we propose that the hybrid propagation step of strand transfer can proceed effectively by branch migration of intact RNA strands up to distances of 32 nt, at least with the help of NC. At distances beyond this, RNase H cleavage of the donor strand is required. With RNase H, a specific invasion site can contribute to transfer to a distance of between 32 and 64 nt. Current results show branch migration occurring efficiently at 17 nt, less efficiently at 24 nt, and nearly nonexistent at 32 nt. Yet, since there is a significant influence of the PCIS at 32 nt we suspect that a combination of proximity effect and branch migration occurs across a hybrid propagation range of 24 nt to a position less than 64 nt. Therefore, we propose that transfers beyond 18 – 24 nt occur either by an acceptor proximity mechanism or by branch migration promoted by cleavages that weaken the donor-cDNA helical stability. This is consistent with previous results demonstrating efficient strand transfer through a proximity mechanism at approximately a distance of 40 nt although branch migration is not fully excluded since there is a possibility for strand exchange at the donor 5'-end which shares 19 nt of homology with the acceptor⁴².

The RNase H and polymerization active sites of the HIV-1 RT are separated by a distance of 16 – 18 nt of the cDNA-donor RNA helix in the RT active site. We envision that the shortest propagation distances for invasion transfer would involve acceptor interaction with an invasion site just as it is leaving the RNase H active site of the RT. Rapid intact-strand branch migration through the substrate cleft of the RT could accomplish transfer over a distance of 17 – 24 nt. We propose this as the minimum distance of hybrid propagation. The influence of a specific invasion site was lost by 32 – 64 nucleotides even in the presence of both NC and RNase H, which promote the hybrid propagation reaction. Therefore, this distance represents the maximum reach of influence of a specific invasion site. Clearly, the many possible sequences and structures of the cDNA-donor and acceptor components of the reaction could affect these distances. This expectation is in general agreement with distributions of terminus transfer sites that we and others have determined with substrates having or lacking hairpin structures^{28; 40; 64; 65; 66}. Nevertheless, our current results point to a relatively narrow range of influence of a specific invasion site for strand transfer. Moreover, our results are fully consistent with the similarity of recombination frequencies in different sections of the viral genome observed during HIV-1 infection.

Materials and Methods

Materials

DNA oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA). The *Taq* DNA Polymerase, dNTP set, *Bam*H I, *Sac* I and *Escherichia coli* Ribonuclease H (RNase

H) were purchased from Invitrogen (San Diego, CA). QIAprep Spin Miniprep Kit was purchased from Qiagen (Valencia, CA). Polynucleotide kinase, T7 RNA polymerase, ribonucleoside triphosphates, and protector RNase Inhibitor were purchased from Roche Applied Science (Indianapolis, IN). The pBluescript I KS (+/-) phagemid vector was purchased from Stratagene (La Jolla, CA). [γ - 32 P] ATP (6000 Ci/mmol) was purchased from PerkinElmer Life Sciences (Boston, MA). Micro bio-spin columns were purchased from Bio-Rad (Hercules, CA). HIV-1 nucleocapsid protein (NCp9) was generously provided by Robert J. Gorelick. The HIV-1 reverse transcriptase (RT) recombinant heterodimer (p66/P51; specific activity 38,000 U/mg) was purified as previously described^{66; 72}. HIV-1 E478Q RT (specific activity 40,000 U/mg) was generously provided by Stuart F. J. Le Grice. Storm PhosphorImager and Image Quant software version 1.2 were from GE Healthcare (Piscataway, NJ).

Generation of RNA and DNA Templates (see Table 1)

Donor DNA (D_{DNA}) was purchased from Integrated DNA Technologies, and subsequently PAGE purified. Donor RNA (D_{RNA}) was created with the use of the pD0 plasmid as a template for PCR amplification along with the primers PCIS 5 (5' – TGGTAAACATTCTTGAGTGCTCGATCTGATGTAAGCTTAGAC – 3') and PCIS 6 (5' – GTAATACGACGACTCACTATAGGGCGAATTGGAGCTCCTCAC – 3'), which introduced base substitutions and the T7 RNA polymerase promoter^{40; 55}. PCR products were gel purified in Ultrapure™ Low Melting Point Agarose and isolated with QIAquick Gel Extraction Kit. Gel-purified PCR products were used as templates for T7 RNA polymerase-catalyzed *in vitro* run-off transcription. Full-length donor RNA products were PAGE purified. Throughout the text, specific nucleotides within the templates are referred to by number (Figure 1, Table 1). The numbering system is based on the final nucleotide at the 3'-end of the donor template, numbered 1, which indicates the first nucleotide after the PCIS. Numbers then increase toward the 5' end of the donor template. Homologous or complementary nucleotides in acceptors or primers, respectively, are numbered in a way that corresponds with the donor template.

The acceptor RNA, A_m , which was used in the strand transfer reactions, was designed to prevent synthesis on the acceptor if the primer exchanged strands from the donor to the acceptor prior to the initiation of synthesis on the donor. To prevent initiation of synthesis on the acceptor, A_m has a 3 nt base substitution at positions 18 – 20, which creates a mismatch with the 3'-terminus of the unextended primers –PCIS₂₀ and +PCIS₂₀. A_m also had a U33A base substitution which was required for synthesis allowing detection of strand exchange products (see *Strand Exchange Assays* below). Generation of A_m required the use of the pA0 plasmid as a template for overlap PCR, with the primers MH47 (5' – GGTTTTCCAGTCACGACG – 3'), PCIS 9 (5' – GGTAAACATTCTTGAGAAGTCGATCTGATGTTAGC – 3'), MH48 (5' – GGAACAAAAGCTGGGTACCG – 3'), and PCIS10 (5' – GCTAACATCAGATCGACTTCTCAAGAATGTTTACC – 3')^{40; 55}. Full-length PCR products were digested with restriction enzymes *Bam*H I and *Sac* I, and subsequently cloned into the pBluescript I KS (+/-) phagemid vector. The pA plasmid was digested with *Bam*H I and used as a template for run-off transcription with T7 RNA polymerase.

A_m was also used in the strand exchange assays to account for the influence of the 3 nt acceptor-primer mismatch at positions 18 – 20 on the kinetics of exchange. The U33A base substitution was employed in the assay to test the 32 nt distance of strand exchange as described in *Strand Exchange Assays* below. Other acceptor templates used in the strand exchange reactions were A_{64} , A_{32} , A_{24} and A_{17} which contain a base substitution at nucleotide positions 65, 33, 25 and 18, respectively. The acceptors were otherwise fully complementary with their respective primers (Table 2). Acceptor RNA, A_{64} , was created using primers PCIS 47 (5' – ACTAATACGACTCACTATAGGGCGAATTGGAGCTCCGATCCTACGAGTATAC

GTCTCACTACACATCGCAGTACAAGTC – 3') and PCIS 28 (5' – CCGGTTCTATAACGGTATGATGGTAAACATTCTTGAGTGCTCGATCTGATGTA AGCTTAGACTTGTACTGCGATGTGTAGTGAG – 3') in a PCR reaction to generate a template for run-off transcription with T7 RNA polymerase. Acceptor RNA, A₁₇, was created using primers PCIS 47 (5' – ACTAATACGACTCACTATAGGGCGAATTGGAGCTCCGATCCTACGAGTATAC GTCTCACTACACATCGCAGTACAAGTC – 3') and PCIS 54 (5' – CCG GTT CTA TAA CGG TAT GAT GGT AAA CAT TCT TGA GAG CTC GAT CTG ATG TAA GCT TAG ACT TGT ACT GCG ATG TGT AGT GAG – 3') in a PCR reaction to generate a template for run-off transcription with T7 RNA polymerase. Acceptor RNA, A₂₄, was created using primers PCIS 47 (5' – ACTAATACGACTCACTATAGGGCGAATTGGAGCTCCGATCCTACGAGTATAC GTCTCACTACACATCGCAGTACAAGTC – 3') and PCIS 55 (5' – CCG GTT CTA TAA CGG TAT GAT GGT AAA CAT TCT TGA GTG CTC GAA CTG ATG TAA GCT TAG ACT TGT ACT GCG ATG TGT AGT GAG – 3') in a PCR reaction to generate a template for run-off transcription with T7 RNA polymerase. Acceptor RNA, A₃₂, was created using primers PCIS 47 (5' – ACTAATACGACTCACTATAGGGCGAATTGGAGCTCCGATCCTACGAGTATACGT CTCCTACTACACATCGCAGTACAAGTC – 3') and PCIS 56 (5' – CCG GTT CTA TAA CGG TAT GAT GGT AAA CAT TCT TGA GTG CTC GAT CTG ATG TTA GCT TAG ACT TGT ACT GCG ATG TGT AGT GAG – 3') in a PCR reaction to generate a template for run-off transcription with T7 RNA polymerase. All full-length acceptor RNA products were PAGE purified.

Preparation of Substrates (see Table 1)

[γ -³²P] ATP (6000 Ci/mmol) was used to radiolabel the following DNA primers: –PCIS₁₇ (5' – TGGTAAACATTCTTGAG – 3'), +PCIS₁₇ (5' – CCGGTTCTATAACGGTATGATGGTAAACATTCTTGAG – 3'), –PCIS₁₈ (5' – TGGTAAACATTCTTGAGA – 3'), +PCIS₁₈ (5' – CCGGTTCTATAACGGTATGATGGTAAACATTCTTGAGA – 3'), –PCIS₂₀ (5' – TGGTAAACATTCTTGAGTGC – 3'), +PCIS₂₀ (5' – CCGGTTCTATAACGGTATGATGGTAAACATTCTTGAGTGC – 3'), –PCIS₂₄ (5' – TGGTAAACATTCTTGAGTGCTCGA – 3'), +PCIS₂₄ (5' – CCGGTTCTATAACGGTATGATGGTAAACATTCTTGAGTGCTCGA – 3'), –PCIS₃₂ (5' – TGGTAAACATTCTTGAGTGCTCGATCTGATGT – 3'), +PCIS₃₂ (5' – CCGGTTCTATAACGGTATGATGGTAAACATTCTTGAGTGCTCGATCTGATGT – 3'), –PCIS₆₄ (5' – TGGTAAACATTCTTGAGTGCTCGATCTGATGTAAGCTTAGACTTGTACTGCGA TGTGTAGTGAG – 3'), +PCIS₆₄ (5' – CCGGTTCTATAACGGTATGATGGTAAACATTCTTGAGTGCTCGATCTGATGTA AGCTTAGACTTGTACTGCGATGTGTAGTGAG – 3'). The primers were separated from unincorporated radionucleotides with Micro bio-spin columns.

Strand Transfer Assays

Radiolabeled primer (0.64 nM), –PCIS₂₀ or +PCIS₂₀, was mixed with donor template (4 nM), either D_{RNA} or D_{DNA} as specified in Figure 2, in 50 mM Tris, pH 8.0, 50 mM KCl, 1 mM DTT, and 1 mM EDTA pH 8.0. The solution was heated at 95° C for 5 min, and then slowly cooled to 37° C. When specified in Figure 2, NC was added to coat the nucleic acid 100 % (1 NC binds 7 nt), followed three minutes later by the addition of RT, either WT or E478Q as specified in Figure 2, to a concentration of 0.16 U/ μ l. The solution was incubated at 37° C for 2 min at which time acceptor RNA A_m (8 nM), MgCl₂ (6 mM), and dNTP (50 μ M) were added simultaneously to start the reaction. Strand transfer reactions were incubated at 37° C for 1, 5,

15, 30 and 60 min, and were stopped with a termination buffer consisting of 90 % formamide, 10 mM EDTA pH 8.0, 0.1 % xylene cyanole, and 0.1 % bromophenol blue.

Strand Transfer Assays with *E. coli* RNase H

Experiments with exogenous *E. coli* RNase H were conducted similarly to the other strand transfer assays except that three minutes after the start of the reaction *E. coli* RNase H was added to the reaction to a final concentration of either 0.02 U/ μ l, or 0.08 U/ μ l. Reactions were incubated for 30 min and terminated with the same buffer used in the other strand transfer assays.

Strand Exchange Assays

Radiolabeled primer (0.64 nM) was mixed with donor template D_{RNA} (4nM), in 50 mM Tris, pH 8.0, 50 mM KCl, 1 mM DTT, and 1 mM EDTA pH 8.0. The solution was heated at 95° C for 5 min, and then slowly cooled to 37° C. When specified in Figures 3 – 6, NC was added to coat the nucleic acid 100 % (1 NC binds 7 nt), which was followed by the addition of either WT or E478Q RT (as noted in Figures 3 – 6) to a concentration of 0.16 U/ μ l. The solution was incubated at 37° C for 2 min at which time acceptor RNA (8 nM), $MgCl_2$ (6 mM), and dNTP (50 μ M) were added simultaneously to start the reaction. Reactions were incubated at 37° C for 1, 3, 5, 15 and 30 min, and were stopped with a termination buffer described in the strand transfer assays.

For strand exchange experiments with the 3 nt mismatch at positions 18 to 20 only acceptor A_m was used in all reactions with primers designed to allow increasing lengths of hybrid propagation. The lengths tested were 17, 24, 32 and 64 nt respectively named –PCIS₁₈, +PCIS₁₈, –PCIS₂₄, +PCIS₂₄, –PCIS₃₂, +PCIS₃₂, –PCIS₆₄ and +PCIS₆₄ (Table 1). Primers either had or were lacking the PCIS as indicated by the – or + signs. Primers –PCIS₁₈ and +PCIS₁₈ had a 1 nt mismatch with the donor, D_{RNA} , which allowed for only 17 base pairs to form between the primer and donor template.

In the strand exchange reactions with complementary primers and acceptors the following combinations of primers and acceptors were used: –PCIS₁₇ or +PCIS₁₇ with A_{17} , –PCIS₂₄ or +PCIS₂₄ with A_{24} , –PCIS₃₂ or +PCIS₃₂ with A_{32} , and –PCIS₆₄ or +PCIS₆₄ with A_{64} . These primers were fully complementary with D_{RNA} and their respective acceptor RNAs (Table 1). To limit primer extension in reactions with the primers –PCIS₁₇, +PCIS₁₇, –PCIS₁₈, +PCIS₁₈, –PCIS₂₄, +PCIS₂₄, –PCIS₂₄ and +PCIS₂₄ only dATP (50 μ M) was used, and reactions with the primers –PCIS₃₂ and +PCIS₃₂ were incubated with only dTTP (50 μ M).

Controls for dNTP incorporation on the acceptor template were conducted in the same manner as the strand exchange reactions except the primer was annealed with acceptor RNA (4 nM) instead of D_{RNA} (4 nM). Also, the reactions were started with the simultaneous addition of $MgCl_2$ (6 mM) and dNTP (50 μ M) but not additional acceptor. Controls for non-template-directed dNTP incorporation were conducted in the same manner as the strand exchange reactions except no acceptor was present in the reaction.

Detection and Analysis

All reaction products were separated by denaturing PAGE, imaged with a Storm PhosphorImager, and quantitated with ImageQuant software version 1.2. Transfer efficiencies (TE) were determined with the equation $TE = (TP / (DE + TP))$, in which TP is the measured intensity of the strand transfer products, and FL is the measured intensity of full-length donor-template extension products. Relative transfer efficiencies (RTE) were determined with the equation $RTE = (TE \text{ of reactions with PCIS} / TE \text{ of reactions without PCIS})$.

Acknowledgments

Special thanks to Dr. Jeffrey DeStefano for critical reading of this manuscript, and many beneficial comments. Thank you to the members of the Bambara laboratory, Dr. Baek Kim, Dr. David H. Mathews and Dr. Lisa M. Demeter for helpful discussions. We thank Dr. Robert J. Gorelick for providing NCp9, and Dr. Stuart F. J. Le Grice for providing E478Q RT. This work was supported by National Institutes of Health grants *Training in HIV Replication and Pathogenesis* 5T32 AI 049815, and GM 049573.

References

1. Boyer JC, Bebenek K, Kunkel TA. Unequal human immunodeficiency virus type 1 reverse transcriptase error rates with RNA and DNA templates. *Proc Natl Acad Sci U S A* 1992;89:6919–23. [PubMed: 1379727]
2. DeStefano JJ, Wu W, Seehra J, McCoy J, Laston D, Albane E, Fay PJ, Bambara RA. Characterization of an RNase H deficient mutant of human immunodeficiency virus-1 reverse transcriptase having an aspartate to asparagine change at position 498. *Biochim Biophys Acta* 1994;1219:380–8. [PubMed: 7522572]
3. Goodenow M, Huet T, Saurin W, Kwok S, Sninsky J, Wain-Hobson S. HIV-1 isolates are rapidly evolving quasispecies: evidence for viral mixtures and preferred nucleotide substitutions. *J Acquir Immune Defic Syndr* 1989;2:344–52. [PubMed: 2754611]
4. Huber HE, McCoy JM, Seehra JS, Richardson CC. Human immunodeficiency virus 1 reverse transcriptase. Template binding, processivity, strand displacement synthesis, and template switching. *J Biol Chem* 1989;264:4669–78. [PubMed: 2466838]
5. Ji JP, Loeb LA. Fidelity of HIV-1 reverse transcriptase copying RNA in vitro. *Biochemistry* 1992;31:954–8. [PubMed: 1370910]
6. Preston BD, Poiesz BJ, Loeb LA. Fidelity of HIV-1 reverse transcriptase. *Science* 1988;242:1168–71. [PubMed: 2460924]
7. Roberts JD, Bebenek K, Kunkel TA. The accuracy of reverse transcriptase from HIV-1. *Science* 1988;242:1171–3. [PubMed: 2460925]
8. Weber J, Grosse F. Fidelity of human immunodeficiency virus type I reverse transcriptase in copying natural DNA. *Nucleic Acids Res* 1989;17:1379–93. [PubMed: 2466238]
9. Burke DS. Recombination in HIV: an important viral evolutionary strategy. *Emerg Infect Dis* 1997;3:253–9. [PubMed: 9284369]
10. Howell RM, Fitzgibbon JE, Noe M, Ren ZJ, Gocke DJ, Schwartzer TA, Dubin DT. In vivo sequence variation of the human immunodeficiency virus type 1 env gene: evidence for recombination among variants found in a single individual. *AIDS Res Hum Retroviruses* 1991;7:869–76. [PubMed: 1760227]
11. Hu WS, Temin HM. Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proc Natl Acad Sci U S A* 1990;87:1556–60. [PubMed: 2304918]
12. Quinones-Mateu ME, Gao Y, Ball SC, Marozsan AJ, Abrahams A, Arts EJ. In vitro intersubtype recombinants of human immunodeficiency virus type 1: comparison to recent and circulating in vivo recombinant forms. *J Virol* 2002;76:9600–13. [PubMed: 12208939]
13. Robertson DL, Hahn BH, Sharp PM. Recombination in AIDS viruses. *J Mol Evol* 1995;40:249–59. [PubMed: 7723052]
14. Temin HM. Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. *Proc Natl Acad Sci U S A* 1993;90:6900–3. [PubMed: 7688465]
15. Vartanian JP, Meyerhans A, Asjo B, Wain-Hobson S. Selection, recombination, and G→A hypermutation of human immunodeficiency virus type 1 genomes. *J Virol* 1991;65:1779–88. [PubMed: 2002543]
16. Hansen J, Schulze T, Moelling K. RNase H activity associated with bacterially expressed reverse transcriptase of human T-cell lymphotropic virus III/lymphadenopathy-associated virus. *J Biol Chem* 1987;262:12393–6. [PubMed: 2442162]
17. Hurwitz J, Leis JP. RNA-dependent DNA polymerase activity of RNA tumor viruses. I. Directing influence of DNA in the reaction. *J Virol* 1972;9:116–29. [PubMed: 4333538]

18. Leis JP, Hurwitz J. RNA-dependent DNA polymerase activity of RNA tumor viruses. II. Directing influence of RNA in the reaction. *J Virol* 1972;9:130–42. [PubMed: 4333539]
19. Starnes MC, Cheng YC. Human immunodeficiency virus reverse transcriptase-associated RNase H activity. *J Biol Chem* 1989;264:7073–7. [PubMed: 2468665]
20. Telesnitsky, A.; Goff, SP. Strong-stop strand transfer during reverse transcription. In: Skalka, a. M.; Goff, SP., editors. *Reverse Transcriptase*. Cold Spring Harbor Laboratory; Cold Spring Harbor: 1993. p. 49-83.
21. Hwang CK, Svarovskaia ES, Pathak VK. Dynamic copy choice: steady state between murine leukemia virus polymerase and polymerase-dependent RNase H activity determines frequency of in vivo template switching. *Proc Natl Acad Sci U S A* 2001;98:12209–14. [PubMed: 11593039]
22. Nikolenko GN, Svarovskaia ES, Delviks KA, Pathak VK. Antiretroviral drug resistance mutations in human immunodeficiency virus type 1 reverse transcriptase increase template-switching frequency. *J Virol* 2004;78:8761–70. [PubMed: 15280484]
23. Svarovskaia ES, Delviks KA, Hwang CK, Pathak VK. Structural determinants of murine leukemia virus reverse transcriptase that affect the frequency of template switching. *J Virol* 2000;74:7171–8. [PubMed: 10888659]
24. DeStefano JJ, Bambara RA, Fay PJ. The mechanism of human immunodeficiency virus reverse transcriptase-catalyzed strand transfer from internal regions of heteropolymeric RNA templates. *J Biol Chem* 1994;269:161–8. [PubMed: 7506252]
25. DeStefano JJ, Mallaber LM, Rodriguez-Rodriguez L, Fay PJ, Bambara RA. Requirements for strand transfer between internal regions of heteropolymer templates by human immunodeficiency virus reverse transcriptase. *J Virol* 1992;66:6370–8. [PubMed: 1383563]
26. DeStefano JJ, Roberts B, Shriner D. The mechanism of retroviral recombination: the role of sequences proximal to the point of strand transfer. *Arch Virol* 1997;142:1797–812. [PubMed: 9672638]
27. Negroni M, Buc H. Copy-choice recombination by reverse transcriptases: reshuffling of genetic markers mediated by RNA chaperones. *Proc Natl Acad Sci U S A* 2000;97:6385–90. [PubMed: 10829081]
28. Roda RH, Balakrishnan M, Kim JK, Roques BP, Fay PJ, Bambara RA. Strand transfer occurs in retroviruses by a pause-initiated two-step mechanism. *J Biol Chem* 2002;277:46900–11. [PubMed: 12370183]
29. DeStefano JJ, Buiser RG, Mallaber LM, Bambara RA, Fay PJ. Human immunodeficiency virus reverse transcriptase displays a partially processive 3' to 5' endonuclease activity. *J Biol Chem* 1991;266:24295–301. [PubMed: 1722202]
30. Furfine ES, Reardon JE. Reverse transcriptase.RNase H from the human immunodeficiency virus. Relationship of the DNA polymerase and RNA hydrolysis activities. *J Biol Chem* 1991;266:406–12. [PubMed: 1702425]
31. Gopalakrishnan V, Peliska JA, Benkovic SJ. Human immunodeficiency virus type 1 reverse transcriptase: spatial and temporal relationship between the polymerase and RNase H activities. *Proc Natl Acad Sci U S A* 1992;89:10763–7. [PubMed: 1279694]
32. Mizrahi V. Analysis of the ribonuclease H activity of HIV-1 reverse transcriptase using RNA.DNA hybrid substrates derived from the gag region of HIV-1. *Biochemistry* 1989;28:9088–94. [PubMed: 2481501]
33. Fuentes GM, Palaniappan C, Fay PJ, Bambara RA. Strand displacement synthesis in the central polypurine tract region of HIV-1 promotes DNA to DNA strand transfer recombination. *J Biol Chem* 1996;271:29605–11. [PubMed: 8939890]
34. Kelleher CD, Champoux JJ. RNA degradation and primer selection by Moloney murine leukemia virus reverse transcriptase contribute to the accuracy of plus strand initiation. *J Biol Chem* 2000;275:13061–70. [PubMed: 10777611]
35. Schultz SJ, Zhang M, Champoux JJ. Recognition of internal cleavage sites by retroviral RNases H. *J Mol Biol* 2004;344:635–52. [PubMed: 15533434]
36. Schultz SJ, Zhang M, Champoux JJ. Sequence, distance, and accessibility are determinants of 5'-end-directed cleavages by retroviral RNases H. *J Biol Chem* 2006;281:1943–55. [PubMed: 16306040]
37. Smith CM, Smith JS, Roth MJ. RNase H requirements for the second strand transfer reaction of human immunodeficiency virus type 1 reverse transcription. *J Virol* 1999;73:6573–81. [PubMed: 10400754]

38. Wisniewski M, Balakrishnan M, Palaniappan C, Fay PJ, Bambara RA. Unique progressive cleavage mechanism of HIV reverse transcriptase RNase H. *Proc Natl Acad Sci U S A* 2000;97:11978–83. [PubMed: 11035788]
39. Wisniewski M, Balakrishnan M, Palaniappan C, Fay PJ, Bambara RA. The sequential mechanism of HIV reverse transcriptase RNase H. *J Biol Chem* 2000;275:37664–71. [PubMed: 10956669]
40. Hanson MN, Balakrishnan M, Roques BP, Bambara RA. Evidence that creation of invasion sites determines the rate of strand transfer mediated by HIV-1 reverse transcriptase. *J Mol Biol* 2006;363:878–90. [PubMed: 16997325]
41. Heath MJ, Destefano JJ. A complementary single-stranded docking site is required for enhancement of strand exchange by human immunodeficiency virus nucleocapsid protein on substrates that model viral recombination. *Biochemistry* 2005;44:3915–25. [PubMed: 15751967]
42. Song M, Basu VP, Hanson MN, Roques BP, Bambara RA. Proximity and branch migration mechanisms in HIV-1 minus strand strong stop DNA transfer. *J Biol Chem* 2008;283:3141–50. [PubMed: 18073206]
43. Levin JG, Guo J, Rouzina I, Musier-Forsyth K. Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism. *Prog Nucleic Acid Res Mol Biol* 2005;80:217–86. [PubMed: 16164976]
44. Peliska JA, Balasubramanian S, Giedroc DP, Benkovic SJ. Recombinant HIV-1 nucleocapsid protein accelerates HIV-1 reverse transcriptase catalyzed DNA strand transfer reactions and modulates RNase H activity. *Biochemistry* 1994;33:13817–23. [PubMed: 7524664]
45. Stoylov SP, Vuilleumier C, Stoylova E, De Rocquigny H, Roques BP, Gerard D, Mely Y. Ordered aggregation of ribonucleic acids by the human immunodeficiency virus type 1 nucleocapsid protein. *Biopolymers* 1997;41:301–12. [PubMed: 9057495]
46. Dib-Hajj F, Khan R, Giedroc DP. Retroviral nucleocapsid proteins possess potent nucleic acid strand renaturation activity. *Protein Sci* 1993;2:231–43. [PubMed: 8443601]
47. Guo J, Wu T, Anderson J, Kane BF, Johnson DG, Gorelick RJ, Henderson LE, Levin JG. Zinc finger structures in the human immunodeficiency virus type 1 nucleocapsid protein facilitate efficient minus- and plus-strand transfer. *J Virol* 2000;74:8980–8. [PubMed: 10982342]
48. You JC, McHenry CS. HIV nucleocapsid protein. Expression in *Escherichia coli*, purification, and characterization. *J Biol Chem* 1993;268:16519–27. [PubMed: 8344933]
49. Beltz H, Piemont E, Schaub E, Ficheux D, Roques B, Darlix JL, Mely Y. Role of the structure of the top half of HIV-1 cTAR DNA on the nucleic acid destabilizing activity of the nucleocapsid protein NCp7. *J Mol Biol* 2004;338:711–23. [PubMed: 15099739]
50. Hong MK, Harbron EJ, O'Connor DB, Guo J, Barbara PF, Levin JG, Musier-Forsyth K. Nucleic acid conformational changes essential for HIV-1 nucleocapsid protein-mediated inhibition of self-priming in minus-strand transfer. *J Mol Biol* 2003;325:1–10. [PubMed: 12473448]
51. Tsuchihashi Z, Brown PO. DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein. *J Virol* 1994;68:5863–70. [PubMed: 8057466]
52. Urbaneja MA, Wu M, Casas-Finet JR, Karpel RL. HIV-1 nucleocapsid protein as a nucleic acid chaperone: spectroscopic study of its helix-destabilizing properties, structural binding specificity, and annealing activity. *J Mol Biol* 2002;318:749–64. [PubMed: 12054820]
53. Williams MC, Rouzina I, Wenner JR, Gorelick RJ, Musier-Forsyth K, Bloomfield VA. Mechanism for nucleic acid chaperone activity of HIV-1 nucleocapsid protein revealed by single molecule stretching. *Proc Natl Acad Sci U S A* 2001;98:6121–6. [PubMed: 11344257]
54. Chen Y, Balakrishnan M, Roques BP, Bambara RA. Acceptor RNA cleavage profile supports an invasion mechanism for HIV-1 minus strand transfer. *J Biol Chem* 2005;280:14443–52. [PubMed: 15657044]
55. Hanson MN, Balakrishnan M, Roques BP, Bambara RA. Effects of donor and acceptor RNA structures on the mechanism of strand transfer by HIV-1 reverse transcriptase. *J Mol Biol* 2005;353:772–87. [PubMed: 16216274]
56. Lapadat-Tapolsky M, De Rocquigny H, Van Gent D, Roques B, Plasterk R, Darlix JL. Interactions between HIV-1 nucleocapsid protein and viral DNA may have important functions in the viral life cycle. *Nucleic Acids Res* 1993;21:831–9. [PubMed: 8383840]

57. Lapadat-Tapolsky M, Pernelle C, Borie C, Darlix JL. Analysis of the nucleic acid annealing activities of nucleocapsid protein from HIV-1. *Nucleic Acids Res* 1995;23:2434–41. [PubMed: 7543198]
58. You JC, McHenry CS. Human immunodeficiency virus nucleocapsid protein accelerates strand transfer of the terminally redundant sequences involved in reverse transcription. *J Biol Chem* 1994;269:31491–5. [PubMed: 7989315]
59. Chen J, Rhodes TD, Hu WS. Comparison of the genetic recombination rates of human immunodeficiency virus type 1 in macrophages and T cells. *J Virol* 2005;79:9337–40. [PubMed: 15994830]
60. Jetzt AE, Yu H, Klarmann GJ, Ron Y, Preston BD, Dougherty JP. High rate of recombination throughout the human immunodeficiency virus type 1 genome. *J Virol* 2000;74:1234–40. [PubMed: 10627533]
61. Levy DN, Aldrovandi GM, Kutsch O, Shaw GM. Dynamics of HIV-1 recombination in its natural target cells. *Proc Natl Acad Sci U S A* 2004;101:4204–9. [PubMed: 15010526]
62. Onafuwa A, An W, Robson ND, Telesnitsky A. Human immunodeficiency virus type 1 genetic recombination is more frequent than that of Moloney murine leukemia virus despite similar template switching rates. *J Virol* 2003;77:4577–87. [PubMed: 12663764]
63. Zhuang J, Jetzt AE, Sun G, Yu H, Klarmann G, Ron Y, Preston BD, Dougherty JP. Human immunodeficiency virus type 1 recombination: rate, fidelity, and putative hot spots. *J Virol* 2002;76:11273–82. [PubMed: 12388687]
64. Balakrishnan M, Roques BP, Fay PJ, Bambara RA. Template dimerization promotes an acceptor invasion-induced transfer mechanism during human immunodeficiency virus type 1 minus-strand synthesis. *J Virol* 2003;77:4710–21. [PubMed: 12663778]
65. Gao L, Balakrishnan M, Roques BP, Bambara RA. Insights into the multiple roles of pausing in HIV-1 reverse transcriptase-promoted strand transfers. *J Biol Chem* 2007;282:6222–31. [PubMed: 17204480]
66. Roda RH, Balakrishnan M, Hanson MN, Wohrl BM, Le Grice SF, Roques BP, Gorelick RJ, Bambara RA. Role of the Reverse Transcriptase, Nucleocapsid Protein, and Template Structure in the Two-step Transfer Mechanism in Retroviral Recombination. *J Biol Chem* 2003;278:31536–46. [PubMed: 12801926]
67. Chen Y, Balakrishnan M, Roques BP, Bambara RA. Steps of the acceptor invasion mechanism for HIV-1 minus strand strong stop transfer. *J Biol Chem* 2003;278:38368–75. [PubMed: 12878597]
68. Derebail SS, DeStefano JJ. Mechanistic analysis of pause site-dependent and -independent recombinogenic strand transfer from structurally diverse regions of the HIV genome. *J Biol Chem* 2004;279:47446–54. [PubMed: 15342633]
69. Moumen A, Polomack L, Unge T, Veron M, Buc H, Negroni M. Evidence for a mechanism of recombination during reverse transcription dependent on the structure of the acceptor RNA. *J Biol Chem* 2003;278:15973–82. [PubMed: 12595540]
70. Basu VP, Song M, Gao L, Rigby ST, Hanson MN, Bambara RA. Strand transfer events during HIV-1 reverse transcription. *Virus Res* 2008;134:19–38. [PubMed: 18279992]
71. Purohit V, Balakrishnan M, Kim B, Bambara RA. Evidence that HIV-1 reverse transcriptase employs the DNA 3' end-directed primary/secondary RNase H cleavage mechanism during synthesis and strand transfer. *J Biol Chem* 2005;280:40534–43. [PubMed: 16221683]
72. Pandey VN, Kaushik N, Rege N, Sarafianos SG, Yadav PN, Modak MJ. Role of methionine 184 of human immunodeficiency virus type-1 reverse transcriptase in the polymerase function and fidelity of DNA synthesis. *Biochemistry* 1996;35:2168–79. [PubMed: 8652558]

Figure 1. Schematic of primers and templates used for PCIS experiments

The precreated invasion site (PCIS) on selected substrates is indicated. Primer names contain either – or + to indicate the presence or absence of the PCIS. The subscript number in the primer name indicates the length of primer that anneals with the donor template. Nucleotide numbering is indicated by dashed vertical lines and numbers. Black arrows represent the various DNA primers. The hatched line represents the donor RNA. The solid gray line represents the acceptor RNA.

Figure 2. PCIS-mediated strand transfer

a. 5'-radiolabeled DNA primers were heat annealed to either RNA or DNA donors. Reactions were initiated with the addition of Mg^{2+} , dNTPs and acceptor RNA, A_m . -PCIS₂₀ indicates experiments with radiolabeled DNA primers lacking the PCIS, and +PCIS₂₀ indicates the presence of the PCIS. The numbers above the templates indicate the length in nt of each segment. b. Effects of the PCIS, NC and RT-RNase H activity on hybrid propagation. Gel images of E478Q RT with an RNA donor template, WT RT with a DNA donor template, WT RT with an RNA donor donor template. c. Graph symbols defined in E478Q RNA are the same for WT DNA and WT RNA. d. *E. coli* RNase H titration in PCIS strand transfer reactions. RNase H is titrated from 0 U/μl to 0.02, and 0.08 U/μl. e. Graph of transfer efficiencies

calculated from PCIS strand transfer reactions. White bars indicate reactions without exogenous *E. coli* RNase H. Light gray bars indicate reactions with 0.02 U/ μ l of *E. coli* RNase H. Dark gray bars indicate reactions with 0.08 U/ μ l of *E. coli* RNase H. f. Graph of the relative transfer efficiency (RTE) of PCIS reactions. RTE is a ratio of transfer efficiencies with and without the PCIS (see *Materials and Methods*).

Figure 3. Acceptor-Primer hybrid propagation through 64 nt of pre-annealed primer-donor hybrid
a. -PCIS₆₄/+PCIS₆₄, D_{RNA}, and acceptor RNAs indicate the primers and templates used. Only dATP was used in the reactions. Donor and acceptor sequences immediately following the primer 3'-terminus are shown, and nonhomologous bases are underlined. The numbers above the templates indicate the nt distances for various segments of the substrates. b. – e. 3 nt mismatch indicates a mismatch between the primer DNA and the acceptor RNA, A_m, at nucleotide positions 18 – 20. Complementary indicates that the primer and acceptor are fully complementary. Dashed lines and white symbols represent transfer efficiency for -PCIS reactions, while solid lines and black symbols represent +PCIS reactions. Squares indicate extension efficiencies with primed-acceptor only. Diamonds indicate strand exchange efficiencies with primed-donor and free-acceptor. Triangles indicate extension efficiencies with primed-donor only. b. Reactions with WT RT, and no NC. c. Reactions with WT RT, and 100 % NC coating. d. Reactions with E478Q RT, and no NC. e. Reactions with E478Q RT, and 100 % NC coating.

Figure 4. Acceptor-Primer hybrid propagation through 32 nt of pre-annealed primer-donor hybrid Only dTTP was used in the reactions. For symbol designations and reaction conditions see Fig. 3 legend.

Figure 5. Acceptor-Primer hybrid propagation through 24 nt of pre-annealed primer-donor hybrid
Only dATP was used in the reactions. For symbol designations and reaction conditions see Fig. 3 legend. Only results from experiments with full complementarity between acceptor and primer are shown.

Figure 6. Acceptor-Primer hybrid propagation through 17 nt of pre-annealed primer-donor hybrid
Only dATP was used in the reactions. The 3 nt mismatch in A_m is italicized, and the $-PCIS_{18}/+PCIS_{18}$ mismatch with the donor is in boldface. For all other symbol designations and reaction conditions see Fig. 3 legend.

Figure 7. Distances of mechanisms and influences of hybrid propagation

Vertical dashed lines and numbers indicate the distances in nucleotides. The numbering begins with the first nucleotide following the invasion site.

Table 2
Fold increase of strand transfer efficiency in PCIS substrates with variables

All results are normalized to -PCIS, 0 % NC experimental results. -PCIS indicates experiments performed with DNA primers lacking the PCIS, while +PCIS indicates experiments with the PCIS.

		-PCIS ₂₀	-PCIS ₂₀	+PCIS ₂₀	+PCIS ₂₀
RT	D _{onor}	0% NC	100% NC	0% NC	100% NC
E478Q	D _{RNA}	1	1.5	6	19
WT	D _{DNA}	1	0.8	3.8	7.5
WT	D _{RNA}	1	1.7	2.1	3.1