Effects of Unpaired Nucleotides within HIV-1 Genomic Secondary Structures on Pausing and Strand Transfer*

Christian Lanciault and James J. Champoux‡

From the Department of Microbiology, University of Washington School of Medicine, Seattle, Washington 98195

Reverse transcriptase-mediated RNA displacement synthesis is required for DNA polymerization through the base-paired stem portions of secondary structures present in retroviral genomes. These regions of RNA duplex often possess single unpaired nucleotides, or "bulges," that disrupt contiguous base pairing. By using well defined secondary structures from the human immunodeficiency virus, type 1 (HIV-1), genome, we demonstrate that removal of these bulges either by deletion or by introducing a complementary base on the opposing strand results in increased pausing at specific positions within the RNA duplex. We also show that the HIV-1 nucleocapsid protein can increase synthesis through the pause sites but not as efficiently as when a bulge residue is present. Finally, we demonstrate that removing a bulge increases the proportion of strand transfer events to an acceptor template that occur prior to complete replication of a donor template secondary structure. Together our data suggest a role for bulge nucleotides in enhancing synthesis through stable secondary structures and reducing strand transfer.

Strand transfer reactions are also essential for retroviral recombination. It is evident that recombination is an important process for generating retroviral genomic diversity and has had a significant impact on the course of the global HIV-1 epidemic (5–12). Inter-subtype and intra-subtype recombination have been documented in many regions, and in some cases the hybrid viruses cause a large proportion of new infections (circulating recombinant forms) (13–22). Homologous recombination in retroviruses is reported to occur at a high rate and is roughly proportional to the length of sequence homology up to a threshold level, after which there appears to be no additional change in the frequency (23–31). Nonhomologous recombination occurs less frequently but also appears to be mediated by short regions of sequence identity (32).

Recombination can take place during reverse transcription at the stage of either minus or plus strand DNA synthesis (32–37). The "displacement assimilation" model was invoked to explain recombination during plus strand synthesis and involves strand transfer between different DNA templates (38–40). Although recent evidence supports this model (28), it appears that the great majority of transfer events actually occur during minus strand synthesis in vivo (33, 41). This observation suggests that the "copy choice" models involving strand transfer of the nascent DNA between two RNA templates is more relevant to understanding the generation of retroviral diversity. There are several variations of the copy choice recombination model, but all are based on four factors as follows: reverse transcriptase processivity, RNase H activity (42–56), RNA secondary structure (57–66), and the viral nucleocapsid protein (NC) (45, 57, 67–70). Similar to the mechanism of the obligate strand transfers in reverse transcription, recombination relies on degradation of the original or "donor" RNA template by the RNase H activity to make regions of the nascent DNA behind the site of polymerization available for base pairing to a second or "acceptor" RNA template. This process is enhanced by the presence of functional NC. Once a portion of the DNA has annealed to the acceptor template, branch migration promotes complete transfer to the acceptor. In some systems, RT pausing on the donor template appears to be the driving force behind strand transfer and allows the RNase H activity to "catch up" to the nascent DNA 3’ end (44). In other systems it is the presence of RNA secondary structure on the acceptor template that dictates the efficiency of strand transfer after RNase H has degraded the donor template (62, 64).

We have shown previously in vitro that HIV-1 RT primer extension reactions through duplex RNA accumulate significant levels of pause products after only limited synthesis. We further demonstrated that the presence of an unpaired nucleotide disrupting RNA base pair contiguity diminished the strength of pause sites at template positions 2 bp on either side of the bulge. The reduction in pausing resulted from a combi-
nation of increased HIV-1 RT processivity, melting of RNA duplex ahead of the DNA primer, and attenuation of branch migration (71). Interestingly, most secondary structures in retroviral genomes possess single or multiple nucleotide bulges in the otherwise duplex stem portions of stem-loop structures. Because some strand transfer is associated with pausing around secondary structures (26, 44, 45), we wanted to determine whether the presence of bulge nucleotides in well defined secondary structures of the HIV-1 genome affect the extent of DNA polymerization through the structures as well as the efficiency of strand transfer along their sequences.

In this work we show that bulge residues can have a significant enhancing effect on DNA synthesis through duplex RNA portions of the HIV-1 TAR element and poly(A) signal. Removing or matching these nucleotides resulted in increased pausing at template positions within the 2-bp window described above. The presence of HIV-1 NC increased primer extension efficiency without affecting the pause pattern; however, simply re-introducing an unpaired nucleotide into the duplex returned the primer extension efficiency to levels observed for extension on the unmodified (wild type) TAR and poly(A) signal sequences. We further demonstrate that removing a bulge nucleotide in the TAR element increases the amount of strand transfer that occurs prior to complete synthesis through the secondary structure. How these results fit into the copy choice recombination models as well as the evolution of viral secondary structures is discussed.

EXPERIMENTAL PROCEDURES

Materials—Reagents were obtained as follows: HIV-1 reverse transcriptase from Worthington; restriction enzymes, T4 polynucleotide kinase, T4 DNA polymerase, T4 DNA ligase, and Vent polymerase for PCR from New England Biolabs; Plasmid pGEM9zf(−) and the TT and SP6 Ribomax transcriptions kits from Promega; [γ-32P]ATP from PerkinElmer Life Sciences; Pfu DNA polymerase for site-directed mutagenesis from Stratagene; and reagents used for RNA solutions from G5Sacc, WT TAR, and G5Lacc, with ScaI. RNA was generated from either a T7 or Sp6 transcription kit. The RNA was denatured, washed twice with 7% formamide, 20 mM EDTA electrophoresis buffer, boiled for 3 min, and purified by denaturing PAGE. The full-length RNA was visualized by UV shadowing, excised, and eluted for at least 18 h in a 1:1 v/v mixture of TE and acidic phenol/chloroform. The mixture was centrifuged, and the aqueous phase was removed. The aqueous phase was further purified once with chloroform and the RNA was precipitated with an equal volume of isopropyl alcohol. The RNA was dissolved in 50 mM Tris, pH 8.0, and EDTA were added back to 10 and 1 mM, respectively, and the volume was brought to 50 μl. Concentration of the RNA was determined by UV absorbance at 260 nm.

Primer Extension Assays—The following DNA oligonucleotides were used as primers: TAR-1 (5′-CACCACACACTTGAAGCCTCAAA-G3′), TAR-2 (5′-GGCAACCTTATGGCTTAAGCAGT3′), TAR-3 (5′-CACACACAGCAGGCACACACTAC3′), or PA-1 (5′-GTTACCA-GAGTCAACAAAGCAGGGGC3′). Each DNA primer was gel-purified and 5′-end-labeled in the presence of [γ-32P]ATP by polynucleotide kinase (Sigma). TAR element and poly(A) templates were 5′-end-labeled DNA primer in the presence of 100 mM Tris, pH 8.0, and 100 mM KCl at 1.05 molar ratio by heating to 95°C for 3 min and cooling to 37°C at a rate of 0.02°C/s using a Thermo-Hybird thermocycler (Savant). DTT, MgCl2, and 7 pmol of HIV-1 RT (Worthington) were added, and the mixture was further incubated at 37°C for 5 min. Primer extensions were then started by the addition of dNTPs. Final concentrations of each were in a total volume of 20 μl as follows: 50 mM Tris, pH 8.0, 50 mM KCl, 10 mM DTT, 5 mM MgCl2, 100 μM of each dNTP, 50 pm primer, 100 nm template, and 700 nm RT. At each time point, 3-μl aliquots were removed and transferred to 12 μl of 96% formamide, 20 mM EDTA. Extension products were separated by denaturing PAGE. Gels were dried and exposed to a PhosphorImager screen. Gel images were scanned using a STORM PhosphorImager and analyzed using ImageQuant software.

Primer Extension Assays in the Presence of HIV-1 Nucleocapsid—Reaction conditions were the same as described under “Primer Extension Assays” except for the following modifications. The primer/template ratio was 1:1 making the final template concentration 50 nm. After addition of the DTT, MgCl2, and 7 pmol of HIV-1 RT and incubation for 5 min at 37°C, 37.5 pmol of HIV-1 NC or 75 pmol of bovine serum albumin (BSA) was added, and the reaction was incubated for an additional 5 min prior to the addition of the dNTPs. The amount of NC added was calculated to provide 100% coating of the RNA based on the assumption that one NC molecule binds 7 nucleotides of RNA (74).

Single Round Primer Extension (‘Trap’) Assay—Labeled primer and RNA template were annealed at a 1:1 ratio (final concentration 50 nm each). In Vitro Strand Transfer—The following DNA oligonucleotides were then added except in the case of the pre-trap in which a mixture of 200 μg of heparin and 30 pmol of an unlabeled DNA/RNA primer/template substrate was added first. Reactions were incubated for 5 min at 37°C. A mixture of the heparin, unlabeled substrate, and dNTPs was then added to each tube to start the reactions except for the pre-trap in which only dNTPs were added. Note that the sequence of the unlabeled DNA/RNA primer/template substrate was different from the labeled pair to prevent any potential strand transfer events.

In Vitro Strand Transfer Assay—Reaction conditions were the same as described under “Primer Extension Assays” except for the following modifications. The primer/donor template ratio was 1:1 (final concentration 50 nm each), and after addition of the DTT, MgCl2, and 7 pmol of HIV-1 RT and incubation for 5 min at 37°C, 37.5 pmol of HIV-1 NC was added, and the reaction was incubated for an additional 5 min. In a separate tube, acceptor RNA was incubated with HIV-1 NC (100% coating) for 5 min. Nucleotides were then added to the tube containing the acceptor RNA/NC mix. This mixture was subsequently added to the reaction tube containing the donor template/HIV-1 RT mix to start the reaction. The final ratio of donor to acceptor RNA template was 1:3.

RESULTS

Effects of Removing Single Nucleotide Bulges from the Stem of the HIV-1 TAR Element on RT Synthesis—We have demonstrated previously (71) that bulge nucleotides interrupting duplex RNA contiguity decrease pausing at template positions surrounding the bulge. The HIV-1 TAR element has a duplex transcription, the products were phenol/chloroform-extracted and precipitated with 2.5 volumes of ethanol. The RNA was collected by centrifugation, washed once with 70% ethanol, and dried. RNA pellets were dissolved in 15 μl of 10 mM Tris, pH 8.0, and 1 mM EDTA (TE) and 45 μl of 96% formamide, 20 mM EDTA electrophoresis buffer, boiled for 3 min, and purified by denaturing PAGE. The full-length RNA was visualized by UV shadowing, excised, and eluted for at least 18 h in a 1:1 v/v mixture of TE and acidic phenol/chloroform. The mixture was centrifuged, and the aqueous phase was removed. The aqueous phase was further purified once with chloroform and the RNA was precipitated with an equal volume of isopropyl alcohol. The RNA was dissolved in 50 mM Tris, pH 8.0, and EDTA were added back to 10 and 1 mM, respectively, and the volume was brought to 50 μl. Concentration of the RNA was determined by UV absorbance at 260 nm.
stem consisting of 24 bp with single nucleotide bulges between the 4th and 5th bp and the 15th and 16th bp. There are also three consecutive bases unpaired between the 20th and 21st bp (Fig. 1A). We first wanted to determine whether removing any of these bulge nucleotides would affect primer extension of a labeled DNA through the TAR secondary structure. Five constructs were designed to compare HIV-1 RT-catalyzed DNA polymerization through the normal TAR structure (WT TAR) with modified structures in which nucleotides complementary to a bulge residue were introduced into the opposite strand of the stem to create a new base pair and thus increase base pair contiguity (Fig. 1A). Each new structure was folded, and the thermodynamic stabilities were calculated by using the algorithm of Zuker (Fig. 1A) (75). RNA corresponding to each construct was transcribed in vitro and gel-purified. A DNA primer was annealed to the RNA so that extension started 25 nucleotides upstream of the beginning of the TAR stem duplex. HIV-1 RT was added to extension reactions, and accumulation of the full-length extension products was monitored over time.

In all cases, an increase in pausing around the former position of the bulge was observed (Fig. 1B). The ÿU16 modification resulted in a small increase in pausing at template position 14, and matching the unpaired triplet further up the stem with ÿU16/ÿAGA22-24 showed a small increase in pausing at 24 (Fig. 1B, lanes 9–12 and 13–16, respectively). It should be noted that in the former case, increased pausing at 14 is within the 2-bp window in which a bulge residue affects pausing (71). This also holds true in the latter case since 24 in the modified structure corresponds to 21 in the WT TAR. Matching the unpaired “C” at the base of the stem had the most dramatic effect on primer extension through the predicted structure producing very strong pause sites at 4, 5, and 6 that persisted throughout the duration of the time course (Fig. 1B, lanes 5–8 and 17–20). In this case, the +6 pause template position corresponds to the +5 pause site observed at early time points in the WT TAR extension reaction (Fig. 1B, lanes 1–4). The percent of primer extended to the end of the template (full-length product) was quantified and plotted versus time (Fig. 1C). The self-priming product was
When compared with synthesis through WT TAR, the $\Upsilon_{16}$ modification had the least effect on primer extension. The doubly modified $\Upsilon_{16}/\Upsilon_{AGA22-24}$ RNA decreased the amount of full-length product formed, but the majority of synthesis reached the end of the template after 20 min. Any construct containing the $G_{5}$ modification, however, significantly reduced complete polymerization through the TAR element with less than 10% of the primer reaching the end of the template after 20 min (Fig. 1C). It should be noted that pausing is typically observed at the base of the TAR hairpin when RT first encounters duplex RNA (56). We also observed pausing at this point, but only in samples taken at times less than 2 min (data not shown). Because the first time point taken for Fig 1B was 2 min, synthesis had already proceeded beyond this pause site and was not visible in the analysis. However, this pause site can also be seen at an earlier time point on a similar substrate in Fig. 5.

Based on the results presented in Fig. 1, it was apparent that a decrease in full-length product formation did not directly correspond with overall stem-loop stability but rather seemed to be related to increased duplex stability at local sites. For example, the $G_{5}$ structure is less thermodynamically stable than the $U_{16}/U_{AGA22-24}$ modified structure ($-42.0$ kcal/mol versus $-51.0$ kcal/mol, respectively) but has a much greater negative effect on full-length product accumulation. This of course is because of the disproportionately strong pausing induced by removing the unpaired C bulge. We therefore wanted to test if the strong pausing was a result of the new C-G base pair or if it was a direct result of removing the bulge. To this end, two additional modifications were made to the TAR element. The first was to replace the C-G base pair in $G_{5}$ to an $\Upsilon_{G5}$ structure and the second was to remove the bulge region (Fig. 2).

**FIG. 2. Comparison of RT synthesis through TAR elements with stem modifications.** A, top portion depicts the structure of the unmodified TAR stem (WT TAR) and the stem structures for the indicated modified TAR elements. Folded structures and $\Delta G_i$ values were calculated using mfold. Unpaired nucleotides are circled, and new base pairs or the loss of the bulge region are boxed. Bottom portion is an 8% denaturing polyacrylamide gel comparing primer extension of TARn-1 through these substrates. Numbering of bands and time points are the same as in Fig. 1B. B, top portion as in A showing structures and $\Delta G_i$ values for substrates with a bulge engineered into the TAR stem pause region. Bottom portion is an 8% denaturing polyacrylamide gel of primer extension products accumulating with time. Numbering of bands and time points are the same as in Fig. 1B.
A-U base pair creating A56–U5, and the second was to delete the unpaired C entirely creating ΔC (Fig. 2A). Additionally, single unpaired nucleotides were introduced into the base of the stem on either strand of the duplex to create new bulges, and in one case a mismatch (Fig. 2B). Primer extensions through these new RNA structures were then performed using primer TARn-1 that annealed to the RNA template directly upstream of the first TAR stem base pair. The position of the labeled primer in relation to the start of RNA base pairing had no effect on the pausing observed in this system (data not shown). Regardless of whether there was a C-G base pair, an A-U base pair, or if the C was deleted, strong pausing was observed at positions +5 and +6 (Fig. 2A). However, when a variety of different bulges or a mismatch was present, full-length product formation was restored to levels observed for synthesis through the unmodified TAR element (Fig. 2B and data not shown).

In a previous report, we demonstrated that HIV-1 RT processivity increases through duplex RNA in the vicinity of a bulge nucleotide (71). To confirm that this is also the case for synthesis through the TAR structures, primer extension assays were repeated in the presence of a heparin/unlabeled substrate trap (see “Experimental Procedures”). A pre-trap control indicated sufficient trap was present to sequester all of the HIV-1 RT and prevent any synthesis (Fig. 3, lanes 1 and 2). Processive synthesis was compared for WT TAR, \( \text{G5, A56-} \text{U5, ΔC,} \) and \( \text{ΔC-} \text{G5} \) (Fig. 3, lanes 3–12). When a bulge residue is present, HIV-1 RT is able to extend a greater proportion of primer to the end of the RNA template in one binding event than in the absence of a bulge. In the case of \( \text{G5, A56-} \text{U5,} \) and \( \text{ΔC,} \) the majority of primer is extended no further than the pause sites observed under distributive conditions.

Effects of HIV-1 NC on Primer Extension through Modified TAR Element Structures—The viral nucleocapsid protein was shown to increase HIV-1 RT synthesis through pause sites and is capable of destabilizing RNA secondary structure (77–82). To test the effects of NC on RT synthesis and pausing, primer extension reactions through WT TAR, \( \text{G5, A56-} \text{U5,} \) and \( \text{ΔC} \) were carried out in the presence of HIV-1 NC and compared with identical reactions containing BSA. Enough NC was added to completely coat the RNA template at a ratio of 1 NC molecule per 7 nucleotides (74) (data shown only for WT TAR and \( \text{G5,} \) Fig. 4A). For all substrates, the percentage of primer extended to the end of the template was quantified for the 20-min time point (the end of the time course), and these results are shown in Fig. 4B. Results from these experiments demonstrate that, in the context of strong pausing at +5 and +6, there is a negative correlation between the overall thermodynamic stability of the TAR stem base and the amount of full-length product formed with time for this series of substrates. Approximately 30% of the primer reaches the end of the template after 20 min for ΔC, whereas only 5% reaches the end for the more stable \( \text{G5.} \) The presence of NC significantly enhances synthesis past the pause sites at +4, +5, and +6 encountered in the modified structures. NC increases synthesis through ΔC ~2.5-fold, through A56–U5 about 3.0-fold, and through \( \text{G5} \) ~5-fold. Despite the different enhancement levels, we observed the same trend in full-length product formation and regional thermodynamic stability as in the absence of NC (Fig. 3B). NC has no apparent effect on synthesis through WT TAR, however, suggesting that the presence of a bulge residue is more effective at preventing pausing than NC. This conclusion is further supported by results of primer extension through TAR substrates engineered with unpaired nucleotides on either strand of the stem duplex in the vicinity of the +4 to +6 pause sites in which the level of full-length product formation was comparable with WT TAR (Fig. 2B).

Removing Bulge Nucleotides Increases Pausing within the Stem Portion of the HIV-1 Poly(A) Signal—To further investigate whether bulge nucleotides facilitate DNA polymerization through RNA secondary structures, the well defined poly(A) signal hairpin was modified to remove unpaired nucleotides and increase RNA base pair contiguity. In vitro primer extension assays performed by Klasens et al. (83) showed that removing the two bulges along the stem of the poly(A) signal resulted in the appearance of a new pause site at template positions between the former bulge sites (Fig. 5A). Based on our previous results, we hypothesized that strong pausing would only be observed if both residues were removed since the template pause positions are close to the 4-bp window for each individual bulge residue. To test this hypothesis, we created four constructs in which the poly(A) signal was either unmodified, PAWT, had only one of the two bulges deleted, PA5U and PAΔC, or had both bulges deleted, PAΔU/ΔC. All of these constructs also possessed the unmodified TAR element downstream of the poly(A) signal sequence. A labeled DNA primer, PA-1, was annealed to the RNAs to leave a 3-base gap between the primer 3’ terminus and the start of the poly(A) stem duplex. As anticipated, the unmodified hairpin exhibited virtually no pausing along the template corresponding to synthesis through the stem portion of the structure (Fig. 5B, lanes 1–6). Deleting either one of the unpaired nucleotides individually resulted in an increase in pausing at template positions between the two bulges only in the 1st min of the time course (Fig. 5B, lanes 7–18). Deleting both bulge residues, however, resulted in significant accumulation of pause products at +10

![Fig. 2](http://www.jbc.org/content/2417/23/2417.full)
and +11 (numbering is in relation to the start of the stem duplex), with the +11 intermediate persisting through the duration of the 20-min time course (Fig. 5B, lanes 19–24). Taken together these data indicate that bulge residues within duplex portions of stem loop structures enhance HIV-1 RT-catalyzed DNA polymerization through retroviral genomic RNA secondary structures in vitro.

Removing Bulge Nucleotides Increases Strand Transfer Prior to Complete Extension through the HIV-1 TAR Element—As discussed above, pausing on RNA templates has been directly correlated with increased strand transfer efficiency (44). Up to this point we have demonstrated that removing bulge residues increases RT pausing within the stem portion of both the TAR element and poly(A) signal secondary structures. We therefore hypothesized that the presence of a bulge residue, by decreasing pausing and enabling greater HIV-1 RT processivity, would actually decrease the amount of strand transfer occurring before complete synthesis through a secondary structure. To test this hypothesis, an in vitro strand transfer system was designed using modified and unmodified TAR sequences as donor templates and two different length acceptor templates (Fig. 6A). First, a labeled DNA primer (TAR2) was annealed to a donor template followed by addition of HIV-1 RT and HIV-1 NC (100% coating). Primer extension on the donor template was then initiated by simultaneous addition of dNTPs and an acceptor template. The acceptor template was also preincubated with NC (100% coating) before addition to the reaction. NC was included in the reactions because it has been shown previously to enhance strand transfer and inhibit the formation of self-priming products that could interfere with strand transfer (45, 57, 67–70, 76).

TAR2 annealed 40 nucleotides upstream of the start of the TAR stem duplex but is unable to anneal to the acceptor templates described below. Donor templates were paired to compare transfer from a template possessing a bulge nucleotide at the base of the TAR stem with an otherwise identical template lacking a bulge in this region. Therefore, WT TAR (bulge) was compared with the ΔC (no bulge) because these two templates are identical in the region of the +4 to +6 pause sites. The ΔG5 (no bulge) substrate has an extra G residue in the template strand in this region, so transfer was compared with the ΔC-ΔG5 (bulge) construct. This system has an advantage for testing the effects of pausing on strand transfer because we can increase or decrease pausing at specific sites without modifying the template sequence itself.

Each pair of donor templates was assayed for strand transfer to one of two different length acceptor templates, a long acceptor (Lacc) or a short acceptor (Sacc) (Fig. 6A). For the WT TAR/Lacc pair, the Lacc template consisted of the 40 nucleotides upstream of the start of TAR2 and lengthened by seven nucleotides at the 5′ end. Therefore, strand transfer can take place along the entire length of the acceptor, even if RT reaches the end of the donor template. When transfer does occur, continued synthesis to the end of the acceptor template results in a product seven nucleotides longer than if synthesis ended at the 5′ end of the donor template. A LacC was designed for the ΔC-ΔG5/ΔG5 pair by modifying ΔC-ΔG5 in the same manner as WT TAR. The Sacc for each donor pair consisted of the 40 nucleotides upstream of the start of TAR stem base pairing plus 28 nucleotides into the TAR element. These 28 nucleotides encompass the proximal strand of the stem (in relation to the direction of reverse transcription) and 4 nucleotides of the loop. In the case of ΔC-ΔG5 and ΔG5, the Sacc possesses an extra G residue in the proximal strand sequence. When Sacc is the acceptor template, a transfer event will only be detected if it occurs before HIV-1 RT has synthesized through the first 28 nucleotides of the TAR stem loop on the donor template. If this condition is met, then extension to the end of the acceptor results in accumulation of a shorter extension product. If RT polymerizes past the 28 nucleotides before transfer, however, this product should not accumulate. It should be noted that if transfer to the Sacc occurs, the full-length transfer product could potentially transfer back to the original donor template. This would result in an underestimation of the amount of strand transfer observed. Because the primer and donor template are annealed at a 1:1 ratio and

![Fig. 4. Comparison of extension through WT and modified TAR structures in the presence or absence of HIV-1 NC.](http://www.jbc.org/)

A. 8% polyacrylamide gel of TARn-1 extension through WT TAR and the ΔG5 variant in the absence (lanes 1–6 and 13–18) or presence of HIV-1 NC (lanes 7–12 and 19–24). Reactions without NC contained BSA in the same buffer. B. bar graph comparing the percentage of extended primer reaching the end of the RNA templates (full-length) after 30 min in the absence (black bars) or presence (gray bars) of HIV-1 NC. Each value is the average of three independent experiments (error bars show ± 1 S.D.).

As we have seen, removing bulge residues from the TAR element increases RT pausing within the stem portion of both the TAR element and poly(A) signal secondary structures. We therefore hypothesized that the presence of a bulge residue, by decreasing RT pausing and enabling greater HIV-1 RT processivity, would actually decrease the amount of strand transfer occurring before complete synthesis through a secondary structure. To test this hypothesis, an in vitro strand transfer system was designed using modified and unmodified TAR sequences as donor templates and two different length acceptor templates (Fig. 6A). First, a labeled DNA primer (TAR2) was annealed to a donor template followed by addition of HIV-1 RT and HIV-1 NC (100% coating). Primer extension on the donor template was then initiated by simultaneous addition of dNTPs and an acceptor template. The acceptor template was also preincubated with NC (100% coating) before addition to the reaction. NC was included in the reactions because it has been shown previously to enhance strand transfer and inhibit the formation of self-priming products that could interfere with strand transfer (45, 57, 67–70, 76).

TAR2 annealed 40 nucleotides upstream of the start of the TAR stem duplex but is unable to anneal to the acceptor templates described below. Donor templates were paired to compare transfer from a template possessing a bulge nucleotide at the base of the TAR stem with an otherwise identical template lacking a bulge in this region. Therefore, WT TAR (bulge) was compared with the ΔC (no bulge) because these two templates are identical in the region of the +4 to +6 pause sites. The ΔG5 (no bulge) substrate has an extra G residue in the template strand in this region, so transfer was compared with the ΔC-ΔG5 (bulge) construct. This system has an advantage for testing the effects of pausing on strand transfer because we can increase or decrease pausing at specific sites without modifying the template sequence itself.

Each pair of donor templates was assayed for strand transfer to one of two different length acceptor templates, a long acceptor (Lacc) or a short acceptor (Sacc) (Fig. 6A). For the WT TAR/Lacc pair, the Lacc template consisted of the 40 nucleotides upstream of the start of TAR2 and lengthened by seven nucleotides at the 5′ end. Therefore, strand transfer can take place along the entire length of the acceptor, even if RT reaches the end of the donor template. When transfer does occur, continued synthesis to the end of the acceptor template results in a product seven nucleotides longer than if synthesis ended at the 5′ end of the donor template. A LacC was designed for the ΔC-ΔG5/ΔG5 pair by modifying ΔC-ΔG5 in the same manner as WT TAR. The Sacc for each donor pair consisted of the 40 nucleotides upstream of the start of TAR stem base pairing plus 28 nucleotides into the TAR element. These 28 nucleotides encompass the proximal strand of the stem (in relation to the direction of reverse transcription) and 4 nucleotides of the loop. In the case of ΔC-ΔG5 and ΔG5, the Sacc possesses an extra G residue in the proximal strand sequence. When Sacc is the acceptor template, a transfer event will only be detected if it occurs before HIV-1 RT has synthesized through the first 28 nucleotides of the TAR stem loop on the donor template. If this condition is met, then extension to the end of the acceptor results in accumulation of a shorter extension product. If RT polymerizes past the 28 nucleotides before transfer, however, this product should not accumulate. It should be noted that if transfer to the Sacc occurs, the full-length transfer product could potentially transfer back to the original donor template. This would result in an underestimation of the amount of strand transfer observed. Because the primer and donor template are annealed at a 1:1 ratio and
there is excess RT such that virtually all of the primer is extended, it is likely that this phenomenon will occur only at a low frequency. Moreover, calculating the transfer efficiency (transfer product/(transfer product/H11001 full-length donor product)) (45) for each reaction indicated that from 20 to 30 min the efficiency did not change significantly (data not shown). If a second strand transfer were occurring, this value would be expected to decrease as transfer products are chased into the full-length donor.

Strand transfer assays were performed for the WT TAR/H9004C and the H9004C-ƒG5/ƒG5 donor pairs using their respective Lacc or Sacc as acceptor template (Fig. 6, B and C, respectively). The percentage of primer extended to either the end of the donor template or the end of the acceptor template was then quantified for the final (30-min) time point of the reaction (Table I). Strand transfer to the Lacc was virtually equivalent for all of the donor templates tested. This result suggests that regardless of whether there is pausing at the 4 to 6 region at the base of the TAR stem, overall transfer to the longer template can still take place efficiently. Strand transfer to the Sacc, however, showed significant differences depending on whether a bulge residue was present or not. For the WT TARΔC pair, there was a 4-fold increase in the amount of transfer product when the C bulge was deleted. Similarly, a 6-fold increase in strand transfer in the absence of a bulge was observed when comparing ΔC-ƒG5 and ƒG5.

**DISCUSSION**

We have demonstrated that removing single nucleotide bulges in the stem of either the TAR element or poly(A) signal, either by deleting nucleotides or inserting matching nucleotides, results in increased pausing within the duplex portion of the hairpin (Figs. 1, 2, and 5). This observation is consistent with our previous work showing HIV-1 RT cannot efficiently synthesize through contiguous duplex RNA (71). The presence of NC appears to increase the proportion of primers extended to the end of the RNA template when a strong pause site is present but not to the same degree as having a bulge residue.

We further demonstrate that increased pausing near the base of the TAR stem correlates with increased strand transfer to acceptor templates that are identical to the donor templates up to the first 28 nucleotides of the TAR secondary structure sequence (Sacc). When an acceptor contains the entire donor sequence plus seven extra nucleotides at the 5’ end (Lacc substrates), the proportion of extended primers that transfer from the donor template is nearly identical whether a bulge is present or not (Table I). The Lacc transfer product used to quantify transfer from WT TARΔC and ΔC-ƒG5/ƒG5 donor pairs does not indicate where along the donor template strand transfer occurs. Whether transfer occurs near the pause sites in donor templates ΔC or ƒG5 or after synthesis through the entire donor template, the end product is identical. This is also
FIG. 6. The effects of unpaired nucleotides in the stem region of the TAR element on strand transfer frequency. A, schematic diagram of strand transfer assay. Gray lines represent RNA, and black lines are DNA. The acceptor templates are shortened at the 3' end to prevent TAR2 from annealing prior to DNA polymerization. The vertical dashed line indicates the start of homology between the donor and acceptor templates. B and C, gel electrophoresis images showing the strand transfer products for different conditions. The table summarizes the results, indicating the presence (+) or absence (-) of unpaired nucleotides in the donor and acceptor templates.
true for the "non-bulge" donor templates WT TAR and ΔC-ΔG5. In the case of the transfer assay in the presence of the short acceptor template, however, we can discern if the strong pausing observed for synthesis along the ΔC and ΔG5 donor templates favors transfer to the acceptor template before polymerization beyond the 28th nucleotide of the TAR element sequence on the donor template when compared with their partner donor templates possessing a bulge nucleotide (and therefore exhibiting little pausing: see Figs. 1, 2, 4, and 6). The formation of the Sacc transfer product (Fig. 6) cannot only be detected if transfer to the Sacc from the respective donor template occurs prior to the incorporation of the 29th nucleotide of the TAR sequence on the donor template. If synthesis on the donor template proceeds beyond this point, then even if transfer occurs additional polymerization cannot take place because there is no additional template sequence. More importantly, in this scenario, the extended primer will be longer than the Sacc transfer product and thus would not be included when calculating the percentage of Sacc transfer product.

Strand transfer using the Sacc cannot be directly compared with transfer to the Lac due to differences in folding of the acceptor templates (data not shown), a parameter shown previously to affect strand transfer (62). Between donor template pairs, however, we can compare how the presence or absence of a bulge nucleotide, and by extension the presence or absence of a strong pause site near the base of the TAR stem, can affect strand transfer to the Sacc. For each donor pair with their respective Sacc (WT TAR/ΔC and ΔC-ΔG5, see Fig. 6), we show that a greater proportion of extended primer ends up as the Sacc transfer product in the absence of an unpaired nucleotide (Fig. 6B, lanes 19–24, and Fig. 6C, lanes 13–18, and Table I) than when an unpaired nucleotide is present (Fig. 6, B, lanes 13–18, and C, lanes 19–24, and Table I). This suggests that in our system increased pausing along duplex RNA, without modifying the portion of donor RNA template sequence that is identical to the acceptor template, results in a significant increase in strand transfer occurring prior to RT completing synthesis of the donor template TAR element secondary structure.

RT processivity, the RNase H activity, viral NC, and RNA secondary structure influence the frequency of copy choice recombination in retroviruses. As RT polymerizes the minus strand DNA, the RNase H activity hydrolyzes the donor template RNA. This frees the nascent DNA for base pairing or "docking" to the acceptor RNA (44, 45, 55). Pausing enhances RNase H-mediated hydrolysis of a donor template, which suggests that pause sites may facilitate the docking process by allowing multiple RNase H cleavages to occur in the absence of continued extension (45). However, it appears that RNase H cleavages occurring during synthesis are sufficient to create conditions favorable for strand transfer in vivo (49). Because RNase H does not rapidly degrade the template RNA in the region 15–20 bases behind the 3’ DNA primer terminus, the docking step likely occurs behind the point of extension (84). Once an acceptor RNA has docked, branch migration promotes the complete transfer of the nascent DNA onto the new template ("locking") (44). The locking step is what determines the point of crossover between the two RNA templates and is influenced by how far the nascent DNA is extended after docking has occurred. Factors such as pausing (44, 45), nucleotide misincorporation (48, 51), non-template base addition (85), the presence of a functional NC (45, 57, 68), and nucleotide availability (43, 46, 53) all determine the rate of polymerization and whether branch migration will catch up to the DNA 3’ terminus.

In this and previous work (71), we demonstrated that HIV-1 RT extension through contiguous RNA duplex is inefficient and results in significant pause product accumulation. Unpaired nucleotides in the vicinity of pause sites, however, relieve pausing and facilitate continued DNA synthesis. The TAR and poly(A) signal hairpins are just two examples of secondary structures present in the viral genome that naturally possess single or multiple nucleotide bulges along the duplex stem portion. Therefore, in vivo, RT may not encounter strong pause sites while performing displacement synthesis through these structures. This suggests that other nonduplex dependent pause sites (i.e. stretches of homo-polymeric nucleotides) would be more likely to create conditions favorable for pausing (64, 86).

If there is a strong pause site present within the RNA duplex portion of a secondary structure, the point of crossover will be observed more frequently near the pause site because the interruption of synthesis provides time for the branch migration phase of the locking process to overtake the polymerization terminus. This is evident from our strand transfer experiments in Fig. 6 and Table I. In the absence of pausing, however, branch migration may eventually overtake the 3’ DNA terminus, but at a location further downstream on the acceptor template. A very recent publication by Derebail and DeStefano (87) first proposed this difference due to pausing when comparing strand transfer within regions of the HIV-1 genome that possess or lack a strong pause site. In conjunction with a role in increasing RT processivity, bulges and other regions of the acceptor template readily accessible for base pairing may also influence the rate of branch migration, which could account for crossover points being observed in loop regions (65, 88).

Our data not only have implications for retroviral strand transfer as discussed above, but they also raise the possibility

| Table I: Comparison of strand transfer to long (Lacc) or short (Sacc) acceptor templates in the presence or absence of a bulge nucleotide on the donor template |
|----------------|----------------|----------------|----------------|----------------|
| Donor template | Transfer in presence of Lac | Transfer in presence of Sacc | | |
| | Full-length donor | Transfer product | Full-length donor | Transfer product |
| | % | | % | |
| WT TAR (+) | 9.6 ± 5.0 | 68.2 ± 4.6 | 71.6 ± 7.3 | 6.1 ± 2.0 |
| ΔC (-) | 6.5 ± 4.2 | 72.8 ± 2.3 | 48.9 ± 10.1 | 25.2 ± 6.6 |
| ΔC-ΔG5 (+) | 9.1 ± 3.8 | 71.6 ± 7.2 | 67.8 ± 0.56 | 7.1 ± 2.6 |
| ΔG5 (-) | 6.8 ± 1.5 | 67.6 ± 8.9 | 37.4 ± 6.3 | 42.9 ± 3.9 |

* (+) indicates presence of unpaired nucleotide; (-) indicates absence of unpaired nucleotide.

b Percent of total extended primer after 30 min ± 1 S.D. Results are from three independent experiments.
Effects of Unpaired Nucleotides on HIV-1 Strand Transfer

that at some point during retroviral evolution there was selection for stem-loop structures with interruptions in the otherwise duplex structure of the stem. Selection for secondary structures that are not too stable or too unstable has been demonstrated for both the TAR element and the poly(A) signal (89, 90). The presence of bulge residues destabilizes these two structures, but the unpaired nucleotides have also been shown to interact with viral and cellular proteins (91–95). Based on what is known about how these interactions affect the viral life cycle, these protein interactions should be considered the dominant selective force that determines secondary structure. Therefore, the destabilizing effects of the bulge residues are secondary to their impact on the overall structure necessary for binding cognate proteins. In contrast, if the assumption is made that these secondary structures existed in some form prior to gaining a protein binding function, then different selective pressures may have existed to maintain the sequence in the retroviral genome. In the context of our work, one of the pressures may have been how well RT could polymerize through secondary structures. Our results suggest that a strong pause site within the duplex stem of a hairpin structure would create conditions favorable for a strand transfer event occurring prior to complete replication of the structure. If this event occurred at a position not involving the secondary structure on the acceptor strand, the stem loop would effectively be deleted, or if the virion were heterozygous for this structure, the sequence would be altered in subsequent progeny virus. We speculate that bulge residues, by stabilizing the presence of secondary structures in the viral genome, influenced the pool of secondary structures available for the development of protein-RNA interactions at later points in retroviral evolution.

Acknowledgments—We thank M. H. Zhang, J. Carey, J. B. Leppard, and B. T. Paulson for helpful discussions and technical expertise. We especially thank S. Schultz and H. Interthal for valuable intellectual input and discussion of data.

REFERENCES

1. Haseltine, W. A., and Baltimore, D. (1976) in *ICN-UCLA Symposium on Molecular and Cellular Biology* (Baltimore, D., Huang, A. S., and Fox, F. C., eds) Vol. 4. pp. 175–213, Academic Press, New York

2. Pelinska, J. A., and Benkovic, S. J. (1992) *Science* 258, 1112–1118

3. Coffin, J. M. (1979) *Proc. Natl. Acad. Sci. U. S. A.* 76, 6289–6293

4. Gallo, F., Robertson, D. L., Morrison, S. G., Hui, H., Craig, S., Decker, J., Fultz, P. N., Girard, M., Shaw, G. M., Hahn, B. H., and Sharp, P. M. (1996) *J. Virol.* 70, 6746–6750

5. Kalish, M. L. (1999) *Science* 283, 295–309

6. Howell, R. M., Fitzgibbon, J. E., Noe, M., Ren, Z. J., Gocke, D. J., Schwartzer, D., and Goudsmit, J. (2003) *AIDS Res. Hum. Retroviruses* 19, 1055–1064

7. Pollakos, G., Abebe, A., Kliphuis, A., De Wit, T. F., Fiesch, B., Tegbaru, B., Moutouh, L., Corbeil, J., and Richman, D. D. (1996) *J. Virol.* 70, 1112–1118

8. Rhodes, T., Wargo, H., and Hu, W. S. (2003) *J. Virol.* 77, 11193–11200

9. Sadowska, E. S., Delviks, K. A., Hwang, C. K., and Pathak, V. K. (2000) *J. Virol.* 74, 1717–1721

10. Takehisa, J., Zekeng, L., Ido, E., Yamaguchi-Kabata, Y., Mbojouleka, I., Harada, Y., Miura, T., Kaptu, L., and Hayami, M. (1999) *J. Virol.* 73, 6810–6820
Effects of Unpaired Nucleotides on HIV-1 Strand Transfer

2423

74. You, J. C., and McHenry, C. S. (1993) J. Biol. Chem. 268, 16519–16527
75. Zuker, M. (2003) Nucleic Acids Res. 31, 3406–3415
76. Driscoll, M. D., Golinski, M. P., and Hughes, S. H. (2001) J. Virol. 75, 672–686
77. Wu, W., Henderson, L. E., Copeland, T. D., Gorelick, R. J., Bosche, W. J., Rein, A., and Levin, J. G. (1996) J. Virol. 70, 7132–7142
78. Urbaneja, M. A., Wu, M., Casas-Finet, J. R., and Karpel, R. L. (2002) J. Mol. Biol. 318, 749–764
79. Ji, X., Klarmann, G. J., and Preston, B. D. (1996) Biochemistry 35, 132–143
80. Golinski, M. P., and Hughes, S. H. (2003) Biochemistry 42, 8153–8162
81. Drummond, J. E., Mounts, P., Gorelick, R. J., Casas-Finet, J. R., Bosche, W. J., Henderson, L. E., Waters, D. J., and Arthur, L. O. (1997) AIDS Res. Hum. Retroviruses 13, 533–543
82. Bernacchi, S., Stoylov, S., Piemont, E., Ficheux, D., Roques, B. P., Darlix, J. L., and Mely, Y. (2002) J. Mol. Biol. 317, 385–399
83. Klasens, B. I., Huthoff, H. T., Das, A. T., Jeeninga, R. E., and Berkhout, B. (1999) Biochim. Biophys. Acta 1444, 355–370
84. DeStefano, J. J., Buiser, R. G., Mallaber, L. M., Bambara, R. A., and Fay, P. J. (1991) J. Biol. Chem. 266, 24295–24301
85. Golinski, M. P., and Hughes, S. H. (2002) Virology 294, 122–134
86. Klarmann, G. J., Schauber, C. A., and Preston, B. D. (1993) J. Biol. Chem. 268, 9793–9802
87. Derebail, S. S., and DeStefano, J. J. (2004) J. Biol. Chem. 279, 47446–47454
88. Galetto, R., Moumen, A., Giacomoni, V., Veron, M., Charnecau, P., and Negroni, M. (2004) J. Biol. Chem. 279, 36625–36632
89. Berkhout, B., Klaver, B., and Das, A. T. (1997) Nucleic Acids Res. 25, 940–947
90. Klaver, B., and Berkhout, B. (1994) EMBO J. 13, 2650–2659
91. Rounseville, M. P., and Kumar, A. (1992) J. Virol. 66, 1688–1694
92. Rounseville, M. P., Lin, H. C., Agbottah, E., Shukla, R. R., Rabson, A. B., and Kumar, A. (1996) Virology 216, 411–417
93. Rothblum, C. J., Jackman, J., Mikovits, J., Shukla, R. R., and Kumar, A. (1995) J. Virol. 69, 5156–5163
94. Klasens, B. I., Thiesen, M., Virtanen, A., and Berkhout, B. (1999) Nucleic Acids Res. 27, 446–454
95. Harrich, D., Mavankal, G., Mette-Snider, A., and Gaynor, R. B. (1995) J. Virol. 69, 4906–4913

Downloaded from http://www.jbc.org/ on July 22, 2018
Effects of Unpaired Nucleotides within HIV-1 Genomic Secondary Structures on Pausing and Strand Transfer
Christian Lanciault and James J. Champoux

J. Biol. Chem. 2005, 280:2413-2423.
doi: 10.1074/jbc.M410718200 originally published online November 12, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M410718200

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 95 references, 56 of which can be accessed free at http://www.jbc.org/content/280/4/2413.full.html#ref-list-1