ACTA ABP BIOCHIMICA POLONICA

Vol. 59, No 4/2012
619–626
on-line at: www.actabp.pl

Regular paper

Low recombination activity of R region located at both ends of the HIV-1 genome

Anna Urbanowicz[1], Anna Kurzyńska-Kokorniak[1], Anna Jankowska[1], Magdalena Alejska[1] and Marek Figlerowicz[1,2]✉

[1]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland; [2]Institute of Computing Science, Poznań University of Technology, Poznań, Poland

Although two strand transfer events are indispensable for the synthesis of double-stranded DNA and establishing HIV-1 infection, the molecular basis of these phenomena is still unclear. The first obligatory template switching event occurs just at the beginning of the virus replication cycle and involves two copies of the 97-nucleotide long R region, located one each at the both ends of the HIV-1 genome (HIV-1 R). Thus, one can expect that the molecular mechanism of this process is similar to the mechanism of homologous recombination which operates in RNA viruses. To verify the above-mentioned hypothesis, we attempted to assess the recombination activity of HIV-1 R. To this end, we tested *in vitro*, how effectively it induces template switching by HIV-1 RT in comparison with another well-characterized sequence supporting frequent homologous crossovers in an unrelated virus (R region derived from *Brome mosaic virus* — BMV R). We also examined if the RNA sequences neighboring HIV-1 R influence its recombination activity. Finally, we tested if HIV-1 R could cause BMV polymerase complex to switch between RNA templates *in vivo*. Overall, our results have revealed a relatively low recombination activity of HIV-1 R as compared to BMV R. This observation suggests that different factors modulate the efficiency of the first obligatory strand transfer in HIV-1 and the homology-driven recombination in RNA viruses.

Key words: strand transfer; template switching; homologous RNA recombination; RNA virus; retrovirus; copy-choice mechanism; R region

Received: 23 May, 2012; revised: 24 September, 2012; accepted: 24 October, 2012; available on-line: 06 November, 2012

INTRODUCTION

In retroviruses, template switching events are inseparable elements of their replication cycle (Holland *et al.,* 1992). At least two strand transfers are required to convert HIV-1 (family *Retroviridae,* subfamily *Orthoretroviridae*, family *Lentivirus*, species *Human immunodeficiency virus 1*) single-stranded genomic RNA into functional double-stranded DNA that is subsequently integrated into the host genome as proviral DNA. The template switching events are also indispensable elements of RNA recombination, which commonly occurs in all types of RNA viruses and retroviruses, and is considered one of the main driving forces of their rapid evolution (Chare & Holmes, 2006; Urbanowicz *et al.,* 2005; Figlerowicz *et al.,* 2003; Borja *et al.,* 1999; Worobey & Holmes, 1999; Nagy & Simon, 1997; Greene & Alison, 1996, 1994; Lai, 1992; Strauss & Strauss, 1988). There are several lines of evidence that template switching by viral polymerase (a copy-choice mechanism) is responsible for the creation of the numerous recombinant forms of HIV-1 circulating worldwide and for the existence of HIV-1 quasi-species (Smyth *et al.*, 2012; Fisher *et al.,* 2010; Powell *et al.,* 2010). The copy-choice model of RNA recombination assumes that two factors are of special importance: (i) the template switching capacity of viral polymerase and (ii) the presence of recombinationally active sequences within donor and acceptor templates. Accordingly, it has been demonstrated that one can affect the crossover frequency as well as the location of recombinant junction sites by introducing specific mutations either in viral polymerase (Alejska *et al.,* 2001; Figlerowicz & Bujarski, 1998; Figlerowicz *et al.,* 1998, 1997; Cornelissen *et al.,* 1997;) or in the donor and acceptor RNA molecules (Alejska *et al.,* 2005, Nagy & Bujarski, 1998, 1997). So far only a few studies comparing the recombination activity of different viruses have been carried out (Chare & Holmes, 2006; Alejska *et al.*, 2005; Shapka & Nagy, 2004; Cheng & Nagy, 2003). Consequently, the question if there are any general rules controlling RNA recombination remains open.

In this paper, we focused on the first obligatory strand transfer involving two copies of HIV-1 R region placed one each at the 5' and 3' ends of viral genomic RNA (Fig. 1A). After the reverse transcription the R regions become a part of 5' and 3' long terminal repeats (5' and 3' LTRs) which enable the integration of the viral double stranded DNA (dsDNA) with the host genome. After the integration, 5' LTR serves as a promoter binding transcription factors and other regulatory proteins indispensable for the expression of HIV-1 genes. Under natural conditions, the first obligatory strand transfer occurs just at the beginning of the HIV-1 replication cycle. This process is initiated when the virus-encoded reverse transcriptase (RT) starts the minus DNA strand synthesis using host tRNALys3 as a primer. The primer binds to the PBS (Fig. 1A) located about 180 nucleotides from the 5'-end of the RNA genome and RT undertakes the synthesis of so-called minus strand strong-stop DNA (Marque *et al.,*1995). After reaching the 5'-end of the genome, RT most probably dissociates from the template and the single stranded DNA (ssDNA) is transferred to the 3'-end of the RNA genome (Muchiri *et al.,* 2011; Berkhout *et al.,*2001). Upon the strand transfer,

✉e-mail: marekf@ibch.poznan.pl

Abbreviations: BMH, branched multiple hairpin; BMV R, R region derived from *Brome mosaic virus;* HIV-1 R, R region located at both ends of HIV-1 genome; LDI, long distance interaction; NC, nucleocapsid protein; PBS, primer binding site; RdRp, RNA dependent RNA polymerase RT, reverse transcriptase; ssDNA, single stranded DNA; TAR, transacting responsive region; UTR, untranslated region.

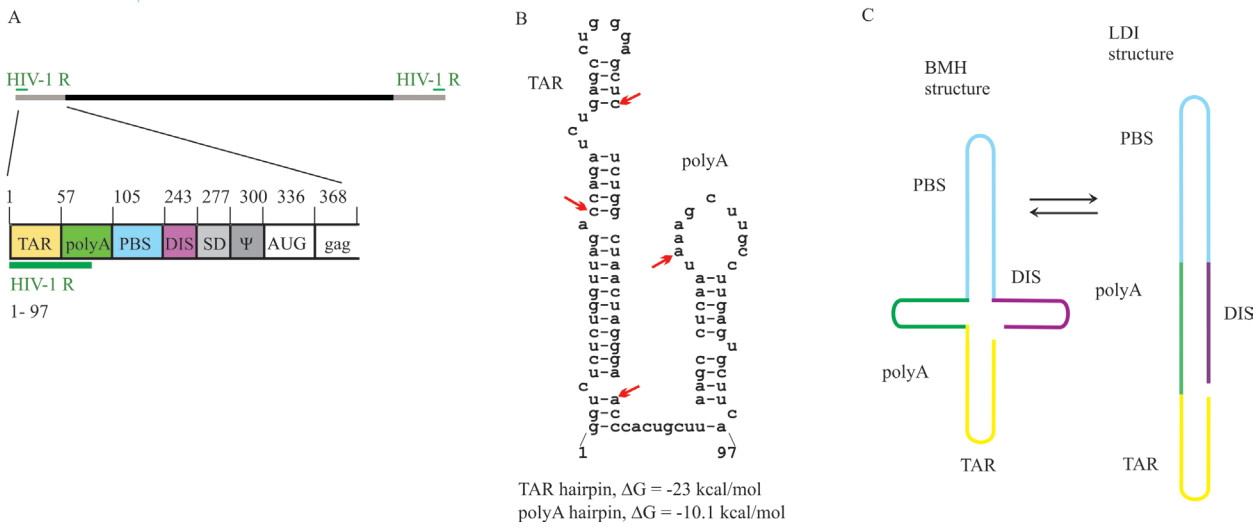


Figure 1. HIV-1 leader sequence and R region.
(A) Organization of the leader sequence and its location in the HIV-1 genome (coding region is marked with a black line, noncoding regions are marked with a grey line). **(B)** HIV-1 R folds in TAR and polyA hairpins. Red arrows point out the marker substitutions in HIV-1Ra. **(C)** Two conformations adopted by HIV-1 leader sequence. More stable structure (long distance interaction, LDI) is formed by long-distance base pairing between polyA and dimerization signal (DIS) regions. A branched multiple structure (BMH), whose formation is assisted by NC protein, exposes the polyA and DIS regions, enabling the formation of RNA dimer.

RT continues the synthesis until the full-length minus DNA strand is completed. The first strand transfer, also called the minus strand transfer, involves hybridization between the ssDNA and the 5'- as well as 3'-terminal R region (Berkhout *et al.,* 2001) (Fig. 1A). In addition to the local homology, the other factors considered to be necessary for the first strand transfer are the DNA polymerase and RNase H activity of HIV-1 RT. Due to the RNase H activity of HIV-1 RT, the copied RNA template is cleaved and removed from the nascent DNA. As a result, the unpaired ssDNA fragment is available for hybridization with the acceptor template (3'-R region). Another important element involved in the first strand transfer is HIV-1-encoded nucleocapsid protein (NC). NC has been proven to strongly facilitate the complementary nucleic acid annealing and strand exchange by destabilizing the secondary structures present in R region (Guo *et al.,*1997) (Fig. 1A, B).

HIV-1 R encloses the TAR hairpin and part of the polyA hairpin (Fig. 1A, B) and is considered a recombinationally active sequence (Moumen *et al.,* 2001). Previously, it has been postulated that the entire HIV-1 leader sequence (nucleotides 1-368) might regulate reverse transcription, including the first obligatory strand transfer by adopting two alternative conformations: LDI and BMH (Berkhout *et al.,* 2002, 2001; Huthoff & Berkhout, 2001) (Fig. 1A, C).

Taking into account a macroscopic description of the first obligatory strand transfer, its high similarity to the homology-driven recombination occurring in RNA viruses can be postulated (Urbanowicz *et al.,* 2005). Both processes involve homologous RNA templates and are mediated by viral polymerases. Earlier it has also been shown that homologous RNA recombination strongly depends on RNA primary and secondary structures. Accordingly, several recombinationally active RNA motifs located in viral genomes have been identified (Alejska *et al.,* 2005; Bruyere *et al.,* 2000; Nagy & Bujarski, 1998, 1997). On the basis of these observations, we attempted here to determine to what extent the effectiveness of the first obligatory strand transfer relies on the recombination activity of HIV-1 R. To this end, we investigated the *in vitro* recombination activity of HIV-1 R and assessed if this activity is affected by the conformational dynamism of the HIV-1 leader sequence. In addition, we tested whether HIV-1 R can efficiently induce *in vivo* recombination if inserted into the genome of a non-related plant RNA virus (family *Bromoviridae*, genus *Bromovirus*, species *Brome mosaic virus* — BMV). Moreover, to assess the template switching capacity of HIV-1 RT, we have analyzed whether this enzyme is capable of producing recombinants in a reaction involving a well-characterized recombinationally active sequence derived from the BMV genome (BMV R). In general, the collected data suggest a relatively low recombination activity of HIV-1 R.

MATERIALS AND METHODS

Materials. pUC19 was purchased from Novagen, restriction enzymes and dNTPs from Promega, Taq DNA polymerase and MMLV reverse transcriptase from MBI Fermentas, RNA was obtained using MEGAshortscript T7 kit (Ambion, Austin, TX, USA). [α^{32}P]dCTP was purchased from ICN. Primers were synthesized by IBB PAS. HIV-1 RT was obtained as previously described by Kurzyńska-Kokorniak *et al,* (2002). pLDI and pBMH plasmids are generous gifts from B. Berkhout (Berkhout *et al.,* 2002).

Plasmids. In order to obtain plasmids containing cDNA of the donor and acceptor RNA templates to be tested *in vitro*, appropriate fragments of the HIV or BMV genome were inserted into pUC19 vector under the control of T7 RNA polymerase promoter. As a result, the following plasmids were generated: pHIV-1Rd containing a 137-nt fragment of HIV-1 5' UTR (97-nt HIV-1 R with 40-nt portion of downstream region U5); pHIV-1Ra containing a 116-nucleotide part of HIV-1 3' UTR cDNA (97-nt HIV-1 R with 19-nt portion of upstream region U3); pBMVR2d containing a BMV-derived sequence located in RNA2 between positions 2640 and 2865 (fragment including BMV R); pBMVR3a containing a BMV-derived sequence located in RNA3 between

positions 1763 and 1961 (fragment including BMV R). To prepare pMatHIVR-RNA3 (containing a full-length cDNA of BMV Mat-HIVR-RNA3), plasmid pMat0-RNA3 (Alejska *et al.,* 2005) was digested with MluI and EcoRV endonucleases and the deleted fragment was replaced with MluI-EcoRV-digested 137-nt cDNA corresponding to the 5' end of HIV-1 genomic RNA. The latter was amplified by PCR from the pHIVR-1d plasmid with primers HIVD5Mlu and HIVD3Eco. Modified pMat0-RNA3 was then digested with SpeI and the deleted fragment was replaced by the SpeI-cut cDNA fragment corresponding to the 116-nucleotide part of 3' HIV-1 UTR cDNA, amplified by PCR from the pHIVRa plasmid with primers HIV A5Spe and HIVA3Spe.

Synthesis of RNA donor and acceptor templates. All RNAs were obtained *via in vitro* transcription from corresponding DNA templates. The DNA templates for transcriptions were synthesized by PCR involving specific primers and above-mentioned plasmids containing selected fragments of the HIV-1 or BMV genomes.

***In vitro* recombination assay.** The template switching capacity of HIV-1 RT was tested in primer extension reactions. They were carried out in a final volume of 40 µl, with 30 pmol of primer, 10 pmol of the donor template and different amounts of the acceptor template. In each set of the experiments involving the HIV-1Rd/HIV-1Ra and BMVR2d/BMVR3a templates, the donor:acceptor molar ratio was as follows: 1:0.5, 1:1, 1:2 and 1:5. In the case of two pairs of templates, BMHd/BMHa and LDId/LDIa, only the 1:5 donor:acceptor ratio was applied. As a control, reactions involving only a donor or acceptor template were carried out. Before each reaction, the primer and donor (primer HIV3-REC for HIV-1Rd, Lys21 for LDId, BMHd and BMHd and BMV3REC for BMVR2d) were denatured at 95ºC in 0.05 M Tris/HCl (pH 8.3), 0.01 M $MgCl_2$, and 0.075 M KCl and slowly cooled (with a cooling rate of 1.0°C/min). When the reaction mixture reached 37ºC, 6 µg of HIV-1 RT was added and the mixture was incubated for further 5 min. Finally, an acceptor, 20 U of ribonuclease inhibitor and a dNTPs mixture (50 µM dATP/dGTP/dTTPmix, 25 µM dCTP, 0.3 µl [α^{32}P]dCTP 3000 Ci/mmol) were added. After a 45-minute incubation at 37°C, the reactions were stopped by the addition of one volume of the urea loading dye (0.25% bromophenol blue and 0.25% xylene cyanoll FF in 7 M urea). For every pair of RNA templates, the experiment was repeated at least three times. Products of the primer extension reactions were denatured by heating at 95°C and fast cooling on ice, separated by electrophoresis in an 8% polyacrylamide denaturing gel (in case of > 100-nt products) or in a 12% polyacrylamide denaturing gel (in case of < 100-nt products) and visualized using a Typhoon phosphorimager. The efficiency of recombination (% of transfer product in total product of primer extension reaction) was quantified by the ImageQuant software. Recombination products were extracted from the gel and amplified with donor-, acceptor- or recombinant-specific primers containing EcoRI and PstI restriction sites. The PCR products were digested with EcoRI and PstI, and ligated into pUC19 vector. The ligation mixture was used to transform DH5α *E. coli* cells and individual clones were sequenced.

***In vivo* recombination assays**. The recombination activity of HIV-1 R was tested in a well-established BMV-based recombination system according to previously described procedure (Alejska *et al.,* 2005). Briefly, BMV genomic RNAs (RNA1, RNA2 and RNA3) were obtained by *in vitro* transcription for which EcoRI-linearized plasmids pB1TP3, pB2TP5, and pMat-HIVR-RNA3 were used. *Chenopodium quinoa* plants (a local lesion host for BMV) were mechanically inoculated with mixtures containing BMV RNA1, RNA2 and Mat-HIVR-RNA3. Two weeks post-inoculation, the number of lesions developed on each inoculated leaf was counted and compared with the number of lesions that appeared after wtBMV infection. To test the recombination activity of the BMV mutant, individual local lesions were excised from the plant leaves and total RNA was extracted separately from each lesion. The isolated RNA was subjected to RT-PCR involving primer 1st and primer 2nd specific for the 3'-portion of RNA3 (the fragment where the tested sequences were located). As a control, RT-PCR reactions involving either the parental Mat-HIVR-RNA3 transcript (positive control) or water (negative control) were carried out. The RT-PCR products were analyzed by electrophoresis in a 1.5% agarose gel, cloned into pUC19 vector and sequenced.

RESULTS

HIV-1 R recombination activity

The ability of HIV-1 R to mediate template switching by HIV-1 RT was tested in reactions involving: 137-nt HIV-1Rd donor RNA (representing 5' end of HIV-1 genome between positions 1 and 137), different amounts of 116-nt HIV-1Ra acceptor RNA (representing 3' end of HIV-1 genome between positions 9516 and 9632) and donor-specific primer HIV3REC (complementary to the 3' end of the donor template between positions 108-132, Fig. 2A–C). In the individual reactions, the donor:acceptor molar ratio was changed as follows: 1:0,5, 1:1, 1:2 and 1:5. Both the donor and acceptor shared the 97-nt HIV-1 R sequence, thus during the primer extension reactions HIV-1 RT could synthesize either a 137-nt ssDNA complementary to the donor template (product $F_{HIV\text{-}1\ R}$) or it could switch to the acceptor, within HIV-1 R, and produce a 178-nt recombinant (product $T_{HIV\text{-}1\ R}$). To estimate the locations of the crossovers, four single-nucleotide marker mutations were introduced into the acceptor template. They were distributed along the entire R region (Fig. 2D).

A polyacrylamide gel electrophoresis of the products formed in the primer extension reactions revealed that the amount of $T_{HIV\text{-}1\ R}$ (putative recombinant) grew proportionally with the increase in the acceptor concentration (donor concentration was constant). The maximal efficiency of $T_{HIV\text{-}1\ R}$ formation was ca. 8% of the total product (Fig. 2B, C). To better characterize the two main products of the primer extension reactions (product $F_{HIV\text{-}1\ R}$ and $T_{HIV\text{-}1\ R}$), they were extracted from the gel and used as templates in PCR involving three pairs of primers: donor-specific, acceptor-specific or recombinant-specific. The PCR with $F_{HIV\text{-}1\ R}$ generated products only for the donor-specific pair of primers, while in the PCR with $T_{HIV\text{-}1\ R}$ only the recombinant-specific pair of primers worked. Both PCR products were cloned and 25 individual clones of each product were sequenced. This confirmed that product $F_{HIV\text{-}1\ R}$ had an identical sequence as the donor template, and product $T_{HIV\text{-}1\ R}$ was recombinant — its 5'-portion derived from the donor and the 3'-portion from the acceptor template. A detailed analysis of the recombined sequences showed that all the crossovers could be classified as precisely homologous. Sixty percent of recombinant junction sites were placed within the 5'-terminal part of R region, after the last

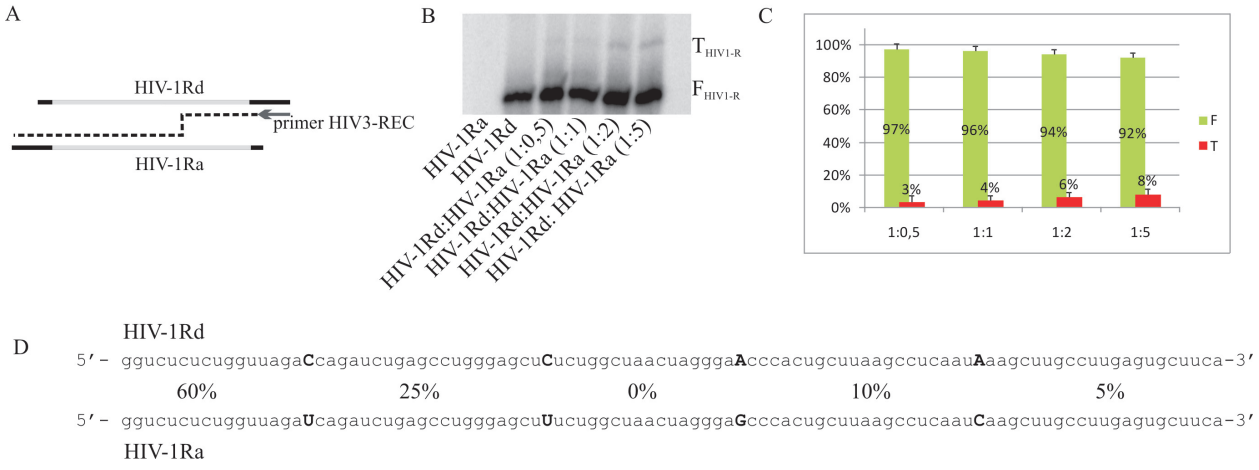


Figure 2. Primer extension reaction with HIV-1 RT and HIV-1Rd/HIV-1Ra.
(A) Schematic description of the HIV-1 RT-based recombination system. Both donor and acceptor templates share HIV-1 R (grey line), the arrow shows the direction of DNA synthesis. **(B)** Example polyacrylamide gel electrophoresis of primer extension reaction products. Reactions were performed with increasing acceptor to donor molar ratio. Full-length products of primer extension reactions are marked $F_{HIV-1\ R}$, while transfer products are marked $T_{HIV-1\ R}$. **(C)** Efficiency of synthesis of strand transfer products. Full-length products of primer extension reactions are marked F, while transfer products are marked T. The efficiency of strand transfer on the graph is shown as the percentage of total reaction product (y axis). The donor:acceptor molar ratio was changed as follows: 1:0.5, 1:1, 1:2, and 1:5 (x axis). **(D)** Distribution of homologous crossovers in HIV-1 R. Marker substitutions are in bold, the distribution of crossovers is shown between the donor and acceptor sequence.

marker mutation, 25% were in the penultimate region, in the middle region no crossovers were observed, while in the first and second regions, respectively, 5% and 10% of crossovers were located (Fig. 2D).

Conformational changes within the HIV-1 leader sequence do not affect its recombination activity

The HIV-1 leader sequence encompasses nucleotides 1–368 of genomic RNA and contains several regulatory elements: the TAR hairpin, the polyA hairpin, PBS, the dimerization initiation site (DIS), the splice donor (SD), the core packaging signal (Ψ) and the hairpin containing the start codon and part of the GAG open reading frame (Fig. 1A). As it was mentioned before, the leader sequence can exist in two alternative conformations (Huthoff & Berkhout, 2001, Fig. 1C). The first more stable LDI is formed by extensive long-distance base pairing between polyA and DIS. The second one, BMH, is adopted in the presence of nucleocapsid protein and consists of several hairpin motifs. The BMH conformation exposes the polyA and the DIS hairpins, enabling the dimerization of HIV-1 genomic RNAs. Considering the fact that conformational changes occurring within the leader sequence can regulate HIV-1 replication (Berkhout *et al.,* 2001; Huthoff & Berkhout, 2001), we attempted to determine if they are also capable of affecting recombination activity of HIV-1 R. Two primer extension reactions were carried out: the one involving a pair of templates with the donor RNA in the LDI conformation, and the second with an analogous pair of templates with the donor adopting the BMH conformation. Both donors (LDId and BMHd) corresponded to nucleotides 1–368 and the acceptors (LDIa and BMHa) to the last 125 nucleotides (nucleotides 9507–9632) of HIV-1 genomic RNA. Since the wild type leader sequence folds in the LDI structure, a few modifications were introduced into BMHd to enforce alternative structure formation, namely: U99 was substituted by C and U91 and C96 were deleted (Berkhout *et al.,* 2002; 2001). Electrophoresis of the LDId and BMHd templates in a native polyacrylamide gel proved that they indeed existed in different conformations (Fig 3A). The *in vitro* recombination assays were done according to a previously described procedure. In all reactions the same donor:acceptor ratio (1:5) was used. The expected length of the product of primer extension on the donor template was 202 nt ($F_{HIV\ BMH}$ or $F_{HIV\ LDI}$), and the length of the recombination product was 235 nt ($T_{HIV\ BMH}$ or $T_{HIV\ LDI}$).

The reaction mixtures were separated by polyacrylamide gel electrophoresis and radioactive products were visualized and quantified with a phosphorimager (Fig. 3B, C). Each band was cut out from the gel, DNA was extracted, amplified and cloned. The sequencing of individual clones revealed that the bands marked as $F_{HIV\ BMH}$ and $F_{HIV\ LDI}$ contained amplified donor templates (BMHd and LDId), while bands marked as $T_{HIV\ BMH}$ and $T_{HIV\ LDI}$ contained recombinants. The band marked as $S_{HIV\ LDI}$ contained the so-called LDId self-priming product. Its formation has been described previously (Discroll *et al.,* 2000). The analysis of $T_{HIV\ BMH}$ and $T_{HIV\ LDI}$ accumulation showed that the conformation adopted by HIV-1 leader sequence did not influence its recombination activity. However, the entire leader sequence facilitated template switching by HIV-1 RT much better than did HIV-1 R alone (20% compared to 8%; Fig. 3D and Fig. 2C, respectively).

Template switching capacity of HIV-1 RT

There are two major factors affecting the efficacy of RNA-RNA recombination: the structure of the donor and acceptor templates and the template switching capacity of viral polymerase. In order to determine which of them was responsible for the relatively low recombination frequency observed during our *in vitro* experiments with HIV-1Ra/HIV-1Rd templates, we used a hybrid recombination system. It was composed of HIV-1 RT and well-characterized, recombinationally active RNA motifs identified earlier in BMV (Nagy & Bujarski, 1995). These BMV-derived motifs are located within 3'UTRs of BMV RNA2 and RNA3 and are called BMV R2 and BMV R3, respectively (Fig. 4A). Both BMV Rs share a 60-nt highly homologous sequence followed by

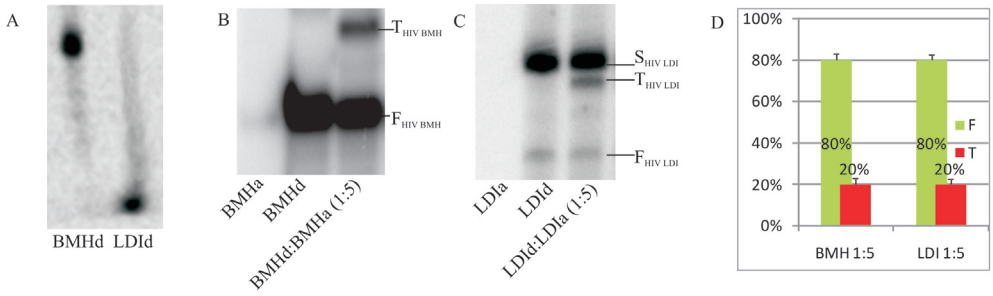
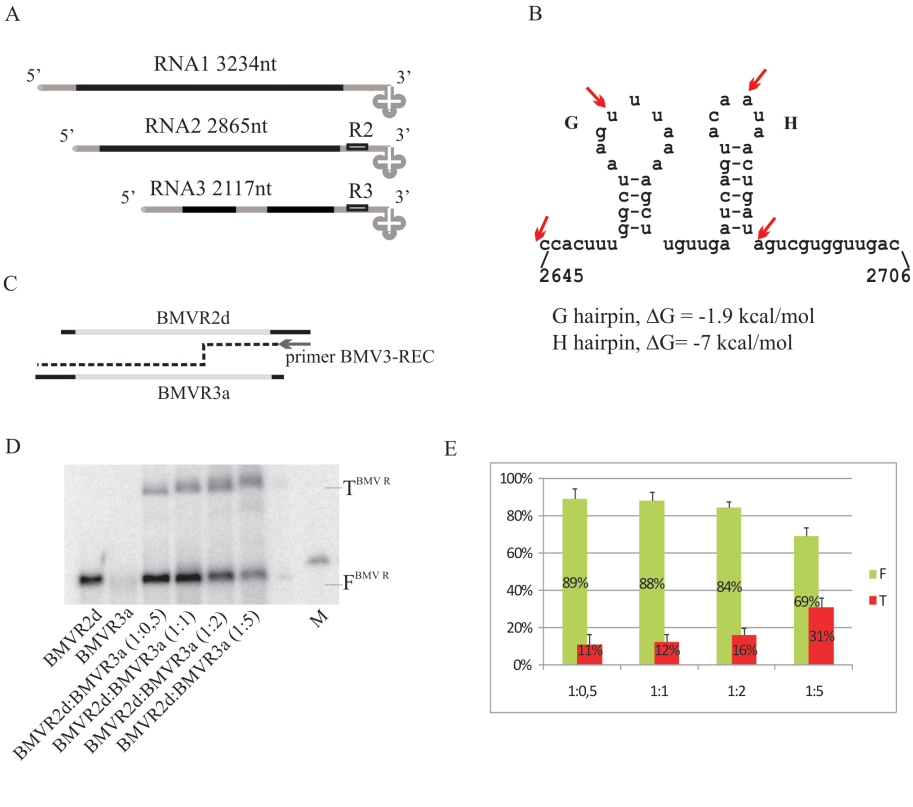


Figure 3. Primer extension reactions with HIV-1 RT and BMHd/BMHa or LDId/LDIa.
(A) The formation of two different structures by LDId and BMHd RNAs was confirmed by their electrophoretic analysis in a native polyacrylamide gel. **(B, C)** Example polyacrylamide gel electrophoresis of products of primer extension reactions with RNA templates adopting BMH **(B)** or LDI **(C)** conformation. Reactions were performed with the donor to acceptor molar ratio of 1:5. Full-length products of the primer extension reactions are marked $F_{HIV\ BMH}$ or $F_{HIV\ LDI}$, while transfer products are marked $T_{HIV\ BMH}$ or $T_{HIV\ LDI}$, self priming products are marked with $S_{HIV\ LDI}$. **(D)** Efficiency of synthesis of strand transfer products. Full-length products of primer extension reactions are marked F, while transfer products are marked T. The efficiency of strand transfer on the graph is shown as the percentage of total reaction product (y axis). Only the 1:5 donor:acceptor ratio was applied.



a 7-nt sequence of reduced homology. Similar to HIV-1 R, a putative secondary structure adopted by BMV R2/3 contains two hairpins called G and H (Nagy & Bujarski, 1995; Fig 4B).

The template switching capacity of HIV-1 RT was tested *in vitro* in primer extension reactions involving two RNA templates: a 225-nt donor RNA (BMVR2d corresponding to nucleotide 2640-2865 of BMV RNA2) and a 214-nt acceptor RNA (BMVR3a encompassing nucleotide 1763-1961 of BMV RNA3, plus a 16-nucleotide sequence added by PCR involving primers T_7 and MB3RNA3) (Fig. 4C–E). The expected recombined product was 364 nt long. Four single-nucleotide mutations present in BMV R3: 1960 C→A, 1946 U→G, 1921 A→G and 1912 A→G (the number of nucleotide corresponds to its location within genomic RNA3) let us estimate the crossover location within the highly homologous 60-nt region (Fig. 4F).

Figure 4. Primer extension reaction with HIV-1 RT and BMVR2d/BMVR3a.
(A) BMV genome. It consists of three RNA molecules, RNA1, RNA2, RNA3, with tRNA-like structures at the 3' ends and the cap structure at the 5' ends (coding regions are marked with black lines, noncoding regions are marked with grey lines, R regions are marked as R2 and R3. **(B)** BMV R adopts secondary structure containing two hairpins G and H. Red arrows point out the marker substitutions in BMVR3a. **(C)** Donor and acceptor templates used in the reaction. Both donor and acceptor templates share BMV R (grey lines), the arrow shows the direction of DNA synthesis. **(D)** Example polyacrylamide gel electrophoresis of primer extension reaction products. Reactions were performed with an increasing acceptor to donor molar ratio. Full-length products of the primer extension reactions are marked $F_{BMV\ R}$, while transfer products are marked $T_{BMV\ R}$. **(E)** Efficiency of synthesis of strand transfer products. Full-length products of primer extension reactions are marked F, while transfer products are marked T. The efficiency of strand transfer on the graph is shown as the percentage of total reaction product (y axis). The donor:acceptor molar ratio was changed as follows: 1:0.5, 1:1, 1:2, and 1:5 (x axis). **(F)** Distribution of homologous crossovers in the BMV R region (60-nt region of high homology is underlined). Marker mutations are in bold, the distribution of crossovers is shown between the donor and acceptor sequence.

The products of the primer extension reactions were separated in a polyacrylamide gel and analyzed with a Phosphorimager (Fig. 4D, E). Two main products of the primer extension reactions (products $F_{BMV\ R}$ and $T_{BMV\ R}$) were extracted from the gel and amplified by PCR involving donor-, acceptor- and recombinant-specific pairs of primers. The PCR with $F_{BMV\ R}$ generated products only for the donor-specific pair of primers, while the PCR with $T_{BMV\ R}$ worked only if the recombinant-specific pair of primers was used. Both PCR products were cloned and sequenced. This confirmed that product $F_{BMV\ R}$ had an identical sequence as the donor template, and product $T_{BMV\ R}$ was recombinant – its 5'-portion derived from the donor and 3'-portion from the acceptor template. The average yield of recombined product synthesis calculated for the reactions with the highest amount of acceptor template (donor:acceptor molar ratio 1:5) reached approximately 30% (Fig. 4D, E). An analysis of the marker mutations present within cloned PCR products demonstrated that all the crossovers were precisely homologous. Sixty percent of the crossovers occurred at the very end of BMV R region within the short sequence (7 nt) showing the reduced level of homology. Twenty-seven percent and 16% of the crossovers took place in the third and fourth region, respectively (Fig. 4F).

HIV-1 R recombination activity in heterologous *in vivo* system

To further characterize the recombination activity of HIV-1 R, we applied a described-earlier BMV-based *in vivo* recombination system (Alejska *et al.,* 2005) (Fig. 5A). The BMV genome consists of three single stranded RNA molecules, RNA1, RNA2 and RNA3. Based on our earlier observations, we created a BMV RNA3 mutant, called Mat-HIVR-RNA3, containing two copies of 97-nt HIV-1 R located within its 3' UTR (5'HIVR and 3'HIVR). The 5' and 3' HIVRs were separated by a spacer sequence of about 350 nt. Additionally, marker substitutions were introduced within 5'HIV-1 R (the same as in HIV1Ra, Fig 2D) to allow us to map the location of recombinant junction sites. We had demonstrated earlier that analogous Mat-RNA3 mutants carrying recombinationally active sequences (located in the same positions as 5'HIVR and 3'HIVR were placed in Mat-HIVR-RNA3) effectively supported homologous recombination in BMV (Alejska *et al.,* 2005). As a result, RNA3 recombinants lacking one homologous region and the whole 350-nt spacer were formed. Because they replicated and accumulated better than the parental RNA3 mutants, the latter were out-competed from the infected plants.

Chenopodium quinoa plants (a local lesion host for BMV) were inoculated with wtRNA1, wtRNA2 and Mat-HIVR-RNA3 transcripts, all the viral components necessary to initiate BMV infection. After two weeks, when the symptoms of infection developed, the lesions formed on every inoculated leaf were counted to estimate the infectivity of the BMV mutant. Forty individual local lesions were excised and total RNA was extracted separately from each of them. Then, the 3'-portion of RNA3 progeny accumulating in examined lesions was selectively amplified by RT-PCR involving RNA3-specific primers 1st and 2nd, flanking the analyzed sequence. The reaction products were separated in a 1.5% agarose gel and their length was determined. The formation of about 800– or 400–500-nt DNA fragments indicated that the lesion contained parental or recombinant RNA3, respectively. The parental-type molecules were detected only in ten local lesions, four lesions contained RNA3 recombinants, while no virus was detected in the remaining 26 local lesions. The RT-PCR-amplified 3'-fragments of RNA3 recombinants (the region where crossovers occur) were cloned and sequenced. As a result, we identified two different types of recombinants. An analysis of their sequences showed that none of them was formed according to the anticipated scenario (HIV-1 R-mediated homologous recombination, see Fig. 5B). Instead, both of them were classified as nonhomologous recombinants and contained a different-size deletion in their 3' UTRs. In the first type the whole spacer sequence was deleted. The second type encompassed the last 120 nucleotides of the spacer sequence, 3'R and 155 nucleotides of 3' UTR.

DISCUSSION

HIV-1 R plays a key role in the first obligatory stand transfer during the conversion of retroviral genomic RNA into ssDNA. Consequently, HIV-1 R has been considered a homologous recombination hot-spot (Moumen *et al.,* 2001). To verify this presumption, we constructed an experimental system allowing us to test the HIV-1 R recombination activity in separation from other factors contributing in the first obligatory strand transfer (Fig. 2A). Surprisingly, our experiments revealed a rela-

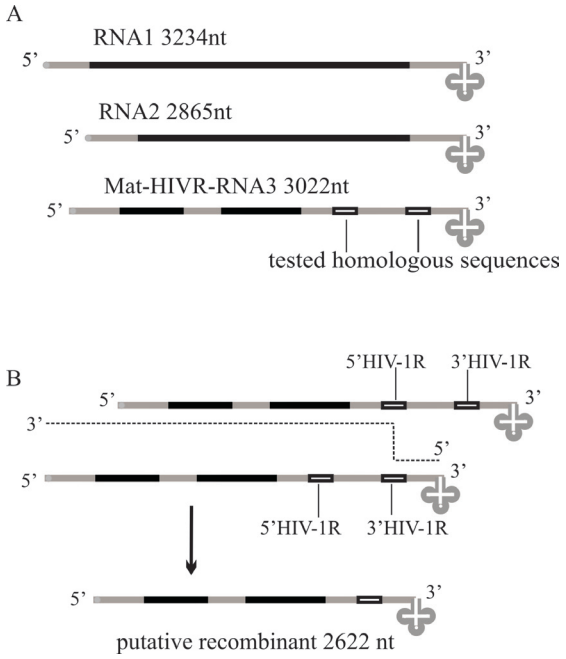


Figure 5. Assessment of HIV-1 R recombination activity in BMV-based *in vivo* system.
(A) In our studies we used the described-earlier Mat-BMV mutant (Alejska *et al.,* 2005). In this mutant specific modifications were introduced only in genomic BMV RNA3. Thus, the system is composed of wtRNA1, wtRNA2 and altered RNA3. Mat-RNA3 has unchanged 5′-noncoding, intergenic and coding regions. The tested homologous sequences (white boxes) are inserted into the 3′-noncoding region and separated by a spacer. As a result, the 3′-noncoding region in Mat-RNA3 is much longer than in wtRNA3. To assess recombination activity of HIV-1 R we constructed MatHIVR-RNA3. This molecule contains 5′ and 3′ HIV1-R as tested sequences. **(B)** The anticipated scenario of HIV-1R-mediated recombination. BMV polymerase initiates nascent RNA strand (dotted line) synthesis at the 3′end of MatHIVR-RNA3. When the polymerase reaches R region it can switch to the homologous region located in the same or other Mat-RNA3 molecule. The resulting RNA3 recombinant lacks one homologous sequence and whole spacer. Because the recombinant replicates and accumulates much better than the parental MatHIVR-RNA3, the latter is easily outcompeted.

tively low recombination activity of the 97-nt HIV-1 R (approximately 8%, Fig. 2B, C), compared to the recombination activity of 60-nt BMV R (approximately 30%, Fig. 4D, E). HIV-1 RT-mediated recombination was the more frequent (up to 20%, Fig. 3B–D) when expanded templates containing the whole HIV-1 leader were used.

We also found that the HIV-1 R sequence was not capable of supporting BMV polymerase-mediated homologous recombination *in vivo*. Certainly, one can question if there is any biological meaning behind this observation. There is no doubt that the BMV-based *in vivo* system cannot be used to study the complex mechanisms underlying the formation of HIV-1 recombinants in human cells. This system, however, seems to be perfect for the assessment of the recombination activity of isolated RNA sequences. It has been designed in such a way that the tested sequence-induced recombination repairs a defective BMV RNA3 molecule (Fig. 5). The repaired RNA3 recombinants are favored by the selective pressure. They replicate and accumulate better than the parental RNA3 and can be easily detected. Accordingly, one can expect that in the employed BMV-based system recombinants are generated provided that the tested sequences are capable of inducing template switching by the viral polymerase.

Earlier it was postulated that the structural polymorphism of the HIV-1 leader may influence the first strand transfer (Berkhout *et al.,*2002, 2001; Huthoff & Berkhout, 2001). However, in our *in vitro* experiments the frequency of recombination was not affected by the leader conformation (BMH or LDI). In natural conditions, the LDI-BMH transformation is enforced by the chaperone activity of viral NC protein (Huthoff & Berkhout, 2001). Recently, it has also been demonstrated that some cellular proteins may bind HIV-1 RNA/DNA and facilitate the first strand transfer (Warrilow *et al.,* 2010). Therefore, one can conclude that several other factors in addition to the local sequence homology and leader structure are involved in this process. The leader region and the nascent DNA might function as a platform enabling the attachment of viral and host proteins modulating the template switch by HIV-1 RT. It seems that the first obligatory strand transfer is a much more complex process than a simple RNA structure-driven RNA recombination.

The data collected so far suggest that stable secondary structures are the major factor that reduces the recombination activity of homologous RNA molecules (Nagy & Bujarski, 1998, 1997, 1995). Two hairpins exist in the both tested sequences, HIV-1 R and BMV R (Fig. 1C, 4B). However, the hairpins present in HIV-1 R are much more stable than those formed within BMV R. The conditions of *in vivo* experiments could additionally increase the stability of the HIV-1 R secondary structure and, consequently, completely inhibit homologous recombination *in planta*. Earlier it was shown that the insertion of a stable stem-loop structure between the G and H hairpins in BMV R negatively influenced the frequency of BMV recombination (Olsthoorn *et al.,* 2002). A positive correlation between the efficiency of template switching by HIV-1 RT and the increasing temperature of primer extension reaction has also been observed. Moreover, mutations destabilizing the TAR and polyA hairpins stimulated strand transfer, while it was inhibited by mutations stabilizing the polyA hairpin (Berkhout *et al.,* 2001; Klavier & Berkhout, 1994). Congenial results were obtained when the recombination activity of two hepatitis virus C (HCV)-derived sequences (highly structured region X and less stable hypervariable region 1) was tested in the BMV-based recombination system described above. The HCV region X adopting a very stable structure did not support recombination and the less structured HVR1 efficiently mediated homologous crossovers (55% of progeny RNA was classified as recombinants (Alejska *et al.*, 2005). All four BMV RNA3 recombinants identified in our *in vivo* assays were products of nonhomologous recombination. This rare process occurs in BMV approximately 10-times less often than homologous recombination (Nagy & Bujarski, 1992). The nonhomologous recombinants were fitter than the parental Mat-HIVR-RNA3 (they probably replicated better than the parental molecules), thus the latter was outcompeted from plant cells. The absence of the virus in more than half of the local lesions (26 out of 40 lesions) also indicated that parental type Mat-HIVR-RNA3 was unable to replicate effectively and accumulated poorly in the plant. Most probably, the HIV-1 R secondary structure was too stable to allow not only the homologous template switching during BMV RNA replication, but it also negatively influenced the replication itself by blocking BMV RdRp movement.

Our *in vitro* assays showed that the HIV-1 RT-mediated crossovers predominantly occurred in 5'-terminal parts of the HIV-1 R and BMV R regions (Fig. 2D and 4F). The majority of the crossovers were located just at the very end of the donor templates. Their location might indicate that HIV-1 RT most effectively cleaves the donor template when it stops after reaching the HIV-1Rd/BMVRd 5' end. In such a situation, when other factors enhancing recombination are absent, template switching by viral polymerase depends on the acceptor capacity to replace the partially digested donor. Earlier it was demonstrated that the latter process can be effectively improved by AU-rich regions, within which donor-nascent strand interactions are weaker. The 5' portion of BMV R is a region of this type (Fig. 4F). In contrast, the 5' portion of HIV-1 R cannot be classified as an AU-rich sequence (Fig. 2D). Consequently, the 5' fragment of BMV R much better supports homologous recombination than the analogous fragment of HIV-1 R.

In the case of HIV-1 R, a number of crossovers were also located in the proximal part of the sequence, namely 5% in the first and 10% in the second part (Fig. 2D). *In vitro* assays carried out by Berkhout's group showed that approximately half of the identified crossovers occurred in the 3' portion of HIV R - before the HIV-1 RT reached the 23rd nucleotide of the HIV R region (Berkhout *et al.,* 2001). This might be a consequence of the reverse transcriptase being stopped by the strong polyA or TAR hairpins (Fig. 1B) and thus allowing extensive RNA cleavage by RNase H (Purohit *et al.,* 2005). RNA secondary structures impede the movement of reverse transcriptase along the template and thus enhance its digestion, since the rate of cleavage is up to 10 times lower than the rate of DNA synthesis (Hanson *et al.,* 2005; Kati *et al.,* 1992). The generated fragments of unpaired ssDNA could then be invaded by the acceptor RNA and the synthesis would be continued using the acceptor RNA (Chen *et al.,* 2005). In the case of the BMV R region, the G and H hairpins are much weaker (Fig. 4B). They could not efficiently slow down HIV-1 RT and enhance the donor template digestion by RNase H which would have increased the possibility of hybridization between a the nascent ssDNA strand and an acceptor RNA template. This fact may explain why crossovers were not found in the 3' portion of the BMV R region.

Here we provided a new piece of evidence that the recombination activity of the HIV-1 R region contrib-

utes little to the first obligatory strand transfer during HIV-1 genomic RNA reverse transcription. Doubtless, several other factors are also involved in the studied phenomenon. Accordingly, it should be considered as a well-synchronized complex process composed of the following events: enforcement of a proper leader structure by NC; RT pausing at a stable hairpin structure or at the HIV-1 R 5' end; RNA template degradation by RNase H; RT stalling and ssDNA exposing; nascent ssDNA strand hybridization with 3'-R region; DNA/RNA hybrid propagation; template replacement. It seems that the relatively low recombination activity of HIV-1 R might in fact be necessary to prevent unwanted template switching events. Numerous factors involved in the first obligatory strand transfer can ensure an effective and strict regulation of this important step of HIV-1 replication.

Acknowledgements

This work was partially supported by the EU Sixth Framework Program through grant no. FP6-037760 and by the Polish Government through grant 6 P04A 038 19.

REFERENCES

Alejska M, Figlerowicz M, Malinowska N, Urbanowicz A, Figlerowicz M (2005) A universal BMV-based RNA recombination system - how to search for general rules in RNA recombination. *Nucl Acids Res* **33**: doi:10.1093/nar/gni105.

Alejska M, Kurzynska-Kokorniak A, Broda M, Kierzek R, Figlerowicz M (2001) How RNA viruses exchange their genetic material. *Acta Biochim Pol* **48**: 391–407.

Berkhout B, Ooms M, Beerens N, Huthoff H, Southern E et al. (2002) In vitro evidence that the untranslated leader of the HIV-1 genome is an RNA checkpoint that regulates multiple functions through conformational changes. *J Biol Chem* **277**: 19967–19975.

Berkhout B, Vastenhouw NL, Klasens BIF, Huthoff H (2001). Structural features in the HIV-1 repeat region facilitate strand transfer during reverse transcription. *RNA* **7**: 1097–1114.

Borja M, Rubio T, Scholthof HB, Jackson AO (1999) Restoration of wild-type virus by double recombination of tombusvirus mutants with a host transgene. *Mol Plant Microbe Interact* **12**: 153–162.

Bruyere A, Wantroba M, Flasinski S, Dzianott A, Bujarski JJ (2000) Frequent homologous recombination events between molecules of one RNA component in a multipartite RNA virus. *J Virol* **74**: 4214–4219.

Chare ER, Holmes EC (2006) A phylogenetic survey of recombination frequency in plant RNA viruses. *Arch Virol* **151**: 933–946.

ChenY, Balakrishnan M, Roques BP, Bambara RA (2005) Acceptor RNA cleavage profile supports an invasion mechanism for HIV-1 minus strand transfer. *J Biol Chem* **280**: 14443–14452.

Cheng C, Nagy PD (2003) Mechanism of RNA recombination in carmo — and tombusviruses: evidence for template switching by the RNA — dependent RNA polymenrase *in vitro*. *J Virol* **77:** 12033–12047.

Cornelissen M, van den Burg R, Zorgdrager F, Lukashov V, Goudsmit J (1997) pol gene diversity of five human immunodeficiency virus type 1 subtypes: evidence for naturally occurring mutations that contribute to drug resistance, limited recombination patterns, and common ancestry for subtypes B and D. *J Virol* **71**: 6348–6358.

Discroll MD, Hughes SH (2000) Human immunodefficiency virus type 1 nucleocapsid protein can prevent self-priming of minus-strand strong stop DNA by promoting the annealing of short oligonucleotides to hairpin sequences. *J Virol* **71**: 2346–2356.

Figlerowicz M, Alejska M, Kurzyńska-Kokorniak A, Figlerowicz M (2003) Genetic variability: the key problem in the prevention and therapy of RNA-based virus infections. *Med Res Rev* **23**: 488–518.

Figlerowicz M, Bujarski JJ (1998) RNA recombination in brome mosaic virus, a model plus stranded RNA virus. *Acta Biochim Pol* **45**: 1–23.

Figlerowicz M, Nagy PD, Bujarski JJ (1997) A mutation in the putative RNA polymerase gene inhibits nonhomologous, but not homologous, genetic recombination in an RNA virus. *Proc Natl Acad Sci USA* **94**: 2073–2078.

Figlerowicz M, Nagy PD, Tang N, Kao CC, Bujarski JJ (1998) Mutations in the N terminus of the brome mosaic virus polymerase affect genetic RNA-RNA recombination. *J Virol* **72**: 9192–9200.

Fisher W, Ganusov VV, Giorgi EE, Hraber PT, Keele BF *et al.* (2010) Transmission of single HIV-1 genomes and dynamics of early immune escape revealed by ultra-deep sequencing. *PLOSone*: **5:** e12303.

Greene AE, Allison RF (1994) Recombination between viral RNA and transgenic plant transcripts. *Science* **263**: 1423–1425.

Greene AE, Allison RF (1996) Deletions in the 3' untranslated region of cowpea chlorotic mottle virus transgene reduce recovery of recombinant viruses in transgenic plants. *Virology* **225**: 231–234.

Guo JH, Henderson LE, Bess J, Kane B, Levin JR (1997) Human immunodeficiency virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-stop DNA. *J Virol* **71**: 5178–5188.

Hanson MN, Balakrishnan M, Roques BP, Bambara RA (2005) Effects of donor and acceptor RNA structures on the mechanism of strand transfer by HIV-1 reverse transcriptase. *J Mol Biol* **353**: 772–787.

Holland JJ, De La Torre JC, Steinhauer DA (1992) RNA virus populations as quasispecies. *Curr Top Microbiol Immunol* **176**: 1–20.

Huthoff H, Berkhout B (2001) Two alternating structures of the HIV-1 leader. *RNA* **7**: 143–157.

Kati WM, Johnson KA, Jerva LF, Anderson KS (1992) Mechanism and fidelity of HIV-1 Reverse transcriptase. *J Biol Chem* **267**: 25988–25997.

Klaver B, Berkhout B (1994) Premature strand transfer by the HIV-1 reverse transcriptase during strong-stop DNA synthesis. *Nucleic Acids Res* **22**: 137–144.

Kurzyńska-Kokorniak A, Jaskólski M, Figlerowicz M (2002) Reverse transcriptase — an enzyme generating the genetic variability of human immunodeficiency virus *Biotechnologia* **1**: 24–34.

Lai MM (1992) RNA recombination in animal and plant viruses. *Microbiol Rev* **56**: 61–79.

Lahser FC, Marsh LE, Hall TC (1993) Contributions of the brome mosaic virus RNA-3 3'-nontranslated region to replication and translation. *J Virol* **67**: 3295–3303.

Marque R, Isel C, Ehresmann C, Ehresmann B (1995) tRNAs as primer of reverse transcriptases. *Biochimie* **77**: 113–124.

Moumen A, Polomack L, Roques B, Buc H, Negroni M (2001). The HIV-1 repeated sequence R as a robust hot-spot for copy-choice recombination. *Nucleic Acids Res* **29**: 3814–3821.

Muchiri JM, Rigby ST, Nguyen LA, Kim B, Bambara RA (2011) HIV-1 Reverse transcriptase dissociates during strand transfer. *J Mol Biol* **412**: 354–364.

Nagy PD, Bujarski JJ. (1992) Genetic recombination in brome mosaic virus: effect of sequence and replication of RNA on accumulation of recombinants. *J Virol* **66**: 6824–6828.

Nagy PD, Bujarski JJ (1995) Efficient system of homologous RNA recombination in brome mosaic virus: sequence and structure requirements and accuracy of crossovers. *J Virol* **69**: 131–140.

Nagy PD, Bujarski JJ (1997) Enigineering of homologous recombination hotspots with AU-rich sequences in brome mosaic virus. *J Virol* **71**: 3799–3810.

Nagy PD., Bujarski JJ (1998) Silencing homologous RNA recombination hot spots with GC-rich sequences in brome mosaic virus. *J Virol* **72**: 1122–1130.

Nagy PD, Simon AE (1997) New insights into the mechanisms of RNA recombination. *Virology* **235**: 1–9.

Olsthoorn RC, Bruyere A, Dzianott A, Bujarski JJ (2002) RNA recombination in brome mosaic virus: effects of strand-specific stem-loop inserts. *J Virol* **76**: 12654-12662.

Powell RLR, Lezeau L, Kinge T, Nyambi PN. (2010) Longitudinal quasispecies analysis of viral variants in HIV type 1 dually infected individuals highlights the importance of sequence identity in viral recombination. *AIDS Res Hum Retroviruses* **26**: 253–264.

Purohit V, Balakrishnan M, Kim B, Bambara RA (2005) 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* **280**: 40534–40543.

Shapka N, Nagy PD (2004) The AU-rich RNA recombination hot spot sequence of brome mosaic virus is functional in tombusviruses: implications for the mechanism of RNA recombination. J *Virol* **78**: 2288–2300.

Smyth RP, Davenporte MP, Maka J (2012) The origin of genetic diversity in HIV-1. *Virus* http://dx.doi.org/10.1016/j.virusres.2012.06.015.

Strauss JH, Strauss EG (1988) Evolution of RNA viruses. *Ann Rev Microbiol* **42**: 657–683.

Urbanowicz A, Alejska M, Formanowicz P, Błażewicz J, Figlerowicz M et al. (2005) Homologous crossovers among molecules of brome mosaic bromovirus RNA1 or RNA2 segments *in vivo*. *J Virol* **79**: 5732–5742.

Warrilow D, Warren K, Harrich D (2010) Strand transfer and elongation of HIV-1 reverse transcription is facilitated by cell factors in vitro. *PLoS One* **5**:e13229.

Worobey M, Holmes E (1999) Evolutionary aspects of recombination in RNA viruses. *J Gen Virol* **80**: 2535–2543.