Downloaded from UvA-DARE, the institutional repository of the University of Amsterdam (UvA)
http://dare.uva.nl/document/84243

File ID	84243
Filename	CHAPTER 7 SUMMARY AND DISCUSSION

SOURCE (OR PART OF THE FOLLOWING SOURCE):

Type	Dissertation
Title	The HIV-1 RNA genome and regulation of reverse transcription and polyadenylation
Author	B.I.F. Klasens
Faculty	Faculty of Medicine
Year	2000
Pages	191

FULL BIBLIOGRAPHIC DETAILS:
http://dare.uva.nl/record/86376

Copyright

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other then for strictly personal, individual use.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)

CHAPTER 7

SUMMARY AND DISCUSSION

GENERAL OVERVIEW

Retroviral genomes contain a terminal redundancy element or repeat (R) region that forms the extreme 5' and 3' ends of this RNA molecule. The human immunodeficiency virus type 1 (HIV-1) encodes a relatively extended R region of 97 nt that folds two adjacent stem-loop structures, the TAR and polyA hairpins. Multiple replicative functions have been attributed to these RNA elements, either as sequence motif or as structured signal, and either as part of the 5' R or the 3' R. The main part of this thesis focuses on an essential function of the HIV-1 polyA hairpin; the regulation of polyadenylation (chapter 2, 3 and 4). Results that support this hypothesis will be summarized and discussed. Although only mentioned briefly, the TAR hairpin may also play a role in the polyadenylation process. A comparison is made for the polyadenylation mechanism of HIV-1 and other complex and simple retroviruses. RNA secondary structure does not only play a role in polyadenylation, but is also shown to also modulate *in vitro* reverse transcription. Stabilization of the polyA hairpin is shown to interfere with elongation process of the RT enzyme (chapter 5), and inhibits the first strand transfer step of the reverse transcription process (chapter 6). These combined results indicate that the thermodynamic stability of the wild-type polyA hairpin is fine-tuned in order to efficiently regulate the polyadenylation process without interfering with other replication processes such as reverse transcription. Initially, we reasoned that the polyA hairpin might also positively influence the reverse transcription process. Another HIV-1 hairpin structure, the dimer linkage structure (DIS), indirectly stimulates reverse transcription. This hairpin structure links the two identical single-stranded RNA transcripts of the viral genome by non-covalent bonds and the close proximity of the two strands has been reported to facilitate the

second strand transfer reaction [311]. For the polyA hairpin no stimulation of reverse transcription was measured. However, for the 5' TAR region we did observe a positive effect on this viral replication step (chapter 6). In order to stress the importance of RNA secondary structure, additional functions of the polyA and TAR hairpin will be summarized and discussed.

POLYADENYLATION STRATEGIES OF RETROVIRUSES

Some retroviruses, such as the mouse mammary tumor virus (MMTV) and the avian sarcoma and leukosis virus (ASLV), contain a relatively small (~20 nt) R region. These viruses encode the AAUAAA polyadenylation (polyA) signal within the U3 region. Transcription starts at the U3/R border of the 5' LTR, proceeds through the viral open reading frames and the 3' LTR into the flanking cellular genome. Thus, the transcription machinery will encounter the AAUAAA hexamer exclusively within the 3' LTR. This passive mechanism ensures that polyadenylation occurs at the 3' end of the viral transcript (illustrated in Fig. 51A).

The human and bovine lymphotropic virus subfamilies (HTLV-I, HTLV-II and BLV) also have the hexamer sequence positioned within the U3 element. However, these viruses contain an R region of ~230 nt. This seems to pose a serious problem, since separation of the AAUAAA sequence and the cleavage site by more than 30 to 40 nt has been reported to lead to a total loss of polyA site usage [64;194;301]. The unusual long R region of these retroviruses contains the Rex Responsive Element (RexRE), a stable RNA secondary structure that is involved in the expression of full-length and singly spliced viral transcripts. The folding of the RexRE into an extended stem-loop structure was proposed to juxtapose the polyA elements [178-180] (illustrated in Fig. 51B). Indeed, mutations that disrupt the predicted structure inhibit the use of the viral RNA processing site [179;180].

Retroviruses like HIV-1 encode the polyA signal within the R region such that this signal is present at both the 5' and 3' ends of the viral transcript (Fig. 51C). This necessitates differential regulation either to repress recognition of the 5' polyA signal or to enhance usage of the 3' signal. Important polyA signals are situated within a stem-loop structure termed the polyA hairpin [117;118].

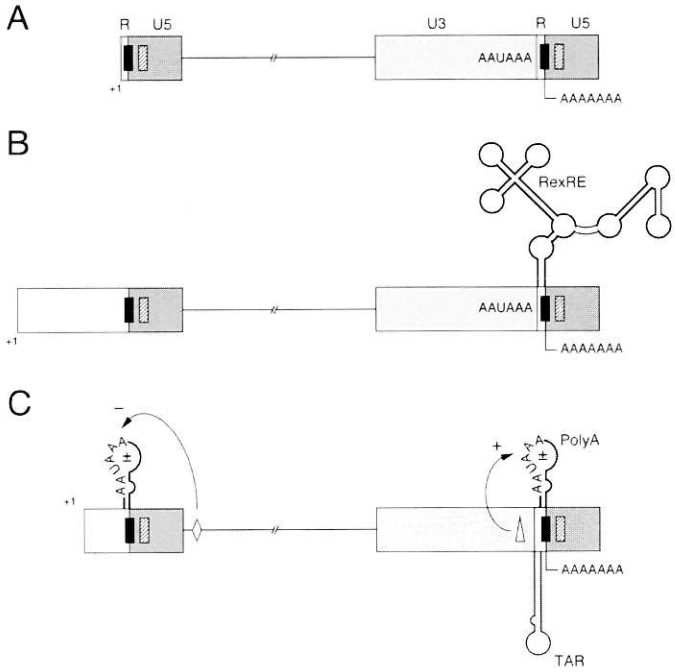

Figure 51. Regulation of polyadenylation for different retroviruses. Retroviral pre-mRNAs contain a R/U5 segment at the 5' end and a complete LTR, that consists of a U3, R and U5 element, at the 3' end. The U3 region is indicated by a grey box, the R region by a white box and the U5 region by a black box in the lower schematics of A, B and C. The LTR varies in size from about 300 to 1200 basepairs for different retroviruses. To simplify the scheme, U3 and U5 regions of similar size were drawn for the different retroviruses. The R region is 16 nt for MMTV (A), 97 nt for HIV-1 (C), and 228 nt for HTLV-I (B), and these differences in size are indicated in the 5' ends of the RNA genome. Folding of the RNA genome into hairpin structures such as TAR and RexRE are illustrated at the 3' ends of the lower schematics. RNA secondary structure can juxtapose polyadenylation signals, illustrated by the differences in length in the R region between the 5' and 3' situation. For MMTV polyadenylation is a passive mechanism that does not require RNA secondary structure (A). For HTLV-I polyadenylation is also a passive mechanism that needs the Rex-Responsive Element (RexRE) to shorten the distance between the AAUAAA hexamer in U3 and the site of cleavage (■) at the R/U5 border (B). For HIV-1 polyadenylation is regulated process that requires the presence of the polyA hairpin at both the 5' and 3' end. The TAR hairpin brings the upstream polyadenylation enhancer element (USE; Δ) close to the AAUAAA hexamer (C). See the text for a detailed description of the differences in regulation of polyadenylation for these retroviruses.

Stabilization of the polyA hairpin was shown to inhibit polyadenylation in transfected cells. The inhibition of polyadenylation for mutants with a stabilized polyA hairpin is the result of reduced binding of the polyadenylation factors to the AAUAAA hexamer in *in vitro* binding assays. The wild-type polyA hairpin partially hinders binding of these factors, but this partial defect can be overcome by the presence of the USE enhancer in the 3'-end specific U3 region. In addition, sequences that are specific for the 5'-end were shown to negatively affect polyadenylation. We propose that the thermodynamic stability of the polyA hairpin is delicately balanced to allow efficient repression at the 5' polyA site, yet full activity of the 3' polyA site.

In an *in vivo* study, viral transcripts with mutant hairpin motifs at either the 5' or 3' end were analyzed for the site of polyadenylation [119]. Premature polyadenylation, e.g. polyadenylation at the 5' site, results in the synthesis of short transcipts of 109 nt. These short RNA products could be detected for the wild-type virus at a low level. Further stabilization of the 5' polyA hairpin resulted in a strong reduction of the synthesis of short RNA products, whereas the amount of premature polyadenylated transcripts was markedly increased for virus constructs with a destabilized polyA hairpin. In the presence of the USE, the situation that is specific for the 3' context, the wild-type polyA hairpin does not interfere with efficient polyadenylation. Stabilization of the 3' polyA hairpin resulted in severe inhibition of polyadenylation at this site. These results confirm that the role of the polyA hairpin is to create a regulatable polyadenylation site. The USE element interacts directly with the 160 kDa subunit of CPSF, the factor responsible for the recognition of the AAUAAA hexamer [82]. The switch of this factor from the USE to the AAUAAA hexamer may be stimulated by a juxtaposition of these two elements by the TAR hairpin. Indeed, this structure has been shown to stabilize polyadenylation complex formation and the polyadenylation efficiency [82;85;177]. Thus, the TAR hairpin is likely to play an additional role in HIV-1 polyadenylation. In the unprocessed primary transcript, the 3' TAR and the 3' polyA hairpins are connected without a single nucleotide between the two stems. This arrangement may result in coaxial stacking, a higher order structure that might bring the USE and the AAUAAA hexamer in close proximity (illustrated in Fig. 51C).

The mechanism of regulation of polyadenylation by the polyA hairpin structure may represent a more common retroviral strategy. Phylogenetic analysis

shows that other retroviruses of the lentivirus and spumavirus groups have a relatively long R region that contains the AAUAAA polyadenylation signal. Similar polyA stem-loop structures could be drawn for these retroviruses [117;119]. There is considerable variation in the thermodynamic stability of these retroviral RNA structures, but stability is merely one of the many parameters that may control the efficiency of these polyadenylation sites. These variables include the actual sequence of the hexamer and perhaps the flanking nucleotide sequences, the presence of enhancer or silencer elements, and the extent of basepairing of these sequences.

THE ROLE OF THE TAR AND POLYA HAIRPIN IN REVERSE TRANSCRIPTION

One obvious function of the retroviral R region is in the process of reverse transcription [41;90;91;93]. Several steps of reverse transcription have been suggested to be either positively or negatively influenced by the structured RNA motifs in the HIV-1 R region. The 5' TAR element has been proposed to stimulate the initiation phase in which the $tRNA^{lys3}$ primer is annealed to the PBS and subsequently extended [115;116]. Furthermore, stable structure in the template RNA can interfere with efficient elongation of the RT enzyme [147;217-220;293]. In this thesis, we showed a direct correlation between the stability of template RNA structure and the extent of RT pausing.

The rules of RT pausing are complex, stops before a basepaired stem was reached were observed on templates with a relative lengthy polyA hairpin. Because the template/primer-bound RT enzyme covers 7 template nucleotides upstream and 22 nucleotides downstream of the cDNA extension point [237], this 'early' stop may reflect the collision of the most frontal RT domain with the basepaired stem. This situation resembles that of the 'toeprinting' technique [238], in which the RT enzyme is blocked by a ribosome bound to the AUG start codon. We also detected pause sites that indicate that the RT enzyme was able to enter the intact stem region of the polyA hairpin up to the site of polymerization. The way of pausing of the RT enzyme may depend in a complex manner on the dimensions of the RNA structure.

In vitro strand transfer assays suggested that stable hairpin structures might stimulate the strand transfer reaction by inducing pausing of the RT enzyme and RNaseH cleavage of the template [267;312]. Stalling of RT during cDNA synthesis was reported to increase the probability of premature template switching [147;266-268]. We obtained different results in a detailed kinetic analysis of reverse transcription. It was demonstrated that cDNA pausing products are eventually chased into full-length ssDNA, and no transfer product was observed until the ssDNA was abundant. Using a sensitive selection-protocol, it was shown that *in vivo* premature strand transfer does occur occasionally [144]. This means that the complete 5' R sequence is inherited in most cases. Furthermore, reverse transcription in virus-infected cells was not significantly affected for mutants with a stabilized polyA hairpin. We therefore suggest that pausing of the RT enzyme is an *in vitro* artifact and that it has no particular role in the strand transfer process.

The 5' RNA template is degraded by RNaseH activity upon reverse transcription. It is therefore not likely that RNA secondary structure in the 5' R region can affect the strand transfer process directly. However, in an *in vitro* reverse transcription assay the TAR region of the 5' R region was shown to stimulate strand transfer. It is tempting to suggest a role for the nucleic acid structure of the ssDNA molecule. The ssDNA molecule is predicted to fold two hairpin structures that are the approximate 'mirror image' of the TAR and polyA hairpins [227]. These two ssDNA hairpins are termed anti-TAR and anti-polyA. It is possible that the capacity of the HIV-1 ssDNA to self-anneal is essential for efficient release from the 5' donor RNA template. Although one may expect that the newly synthesized ssDNA will not be in a duplex with the donor RNA template because of removal of the latter by the RNaseH activity of the RT enzyme, previous reports have indicated that the template RNA is degraded incompletely during reverse transcription [91;236;281;287;290]. RNaseH cleavage occurs infrequently, leaving RNA fragments of considerable length (e.g. 15 to 100 nt). Many of these RNA fragments will not dissociate spontaneously from the newly synthesized ssDNA, thus posing a problem for strand transfer. Another possible function of ssDNA structure is in the subsequent step of strand transfer, in which the ssDNA anneals to the acceptor RNA. Because the RNA and ssDNA structure are nearly mirror-images of each other, there is complete complementarity between the loop regions of the TAR RNA hairpin and the anti-TAR DNA structure, as well as

between the polyA RNA hairpin and the anti-polyA DNA structure. Thus, basepairing interactions between the loops may represent the initial contact, a mechanism that is very similar to the 'loop-loop kissing' interaction during HIV-1 RNA dimerization [127].

In order to anneal to the 3' R RNA template, subsequent melting of the ssDNA secondary structure is required. These basepair rearrangements might be a rate-limiting step in the reverse transcription process. Folding of the TAR and polyA hairpin of the 3' R RNA template was shown to interfere with strand transfer. Thus, folding of the nucleic acid strands may also negatively affect reverse transcription. The overall positive effect of the anti-TAR hairpin in the ssDNA molecule indicates that melting of the hairpin structures is causing a minor problem. This may be explained by the activity of the NC protein that catalyzes basepair rearrangements such as opening of hairpin structures and formation of cDNA-RNA duplexes.

Retroviruses that contain an R region of only ~20 nt, such as MMTV and ASLV, are not supposed to fold extensive inhibitory RNA secondary structure motifs. Therefore, these viruses may not need the stimulating effect of the folding of the ssDNA molecule. In this case, the transfer of the ssDNA molecule from the 5' R to the 3' R is most likely driven solely by the sequence complementarity between the ssDNA and the 3' R. For retroviruses with a longer R region, such as HIV-1, reverse transcription can be affected by multiple factors. The strand transfer process is a complex interplay of elements including NC activity and nucleic acid folding. The latter element can have either a positive or negative effect, dependent whether it involves the 5' or 3' R template. Elucidation of the strand transfer reaction is further complicated by RT pausing in *in vitro* assays. Another *in vitro* artefact that has not been mentioned previously is that annealing of the ssDNA to the donor RNA is preferred over productive binding to the acceptor RNA [227]. This effect is most likely caused by a higher basepair complementarity between the ssDNA molecule and the donor template than between the ssDNA molecule and the acceptor template. This problem can be partially resolved by reducing the amount of donor RNA in the test tube, but it remains questionable whether the strand transfer process can be unraveled in further detail in an *in vitro* reverse transcription assay.

In the context of a replicating virus, synthesis of the viral full-length dsDNA is a far more intricate mechanism. This multistep process that takes place in a nucleoprotein complex that includes different viral and cellular factors. A detailed

understanding of the molecular mechanisms involved in the formation of infectious virus and the structure and composition of the reverse transcription complex is needed to elucidate the process of reverse transcription and the strand transfer steps.

ADDITIONAL REPLICATIVE FUNCTIONS OF THE TAR AND POLYA HAIRPIN

The regulation of polyadenylation and stimulation of reverse transcription by the polyA and TAR RNA hairpins and the anti-TAR cDNA hairpin represent examples that reinforce the idea that critical viral replicative functions may depend on nucleic acid secondary structure. The TAR and polyA hairpins have additional replicative functions. It is generally known that TAR serves as the binding site for the viral tat protein and cellular cofactors in the process of transcriptional activation [18;19]. The 5' TAR motif forms the extreme 5' end of all HIV-1 mRNAs, and has also been suggested to influence translation [313-317]. This RNA structure inhibits translation when introduced into the 5' end of a reporter mRNA. Inhibition is probably due to a combination of factors, including the inability of the translation initiation complex to unwind secondary structure and the inaccessibility of the cap by the folding of RNA for factors involved in initiation of translation [313]. The TAR hairpin also activates the double-stranded RNA-dependent kinase PKR, a protein that inhibits initiation of translation [314-316;318]. Furtermore, the TAR hairpin has been demonstrated to be involved in RNA dimerization in HIV-2 [319]. In HIV-1, electronic microscopic studies are consistent with the involvement of 5' R sequences in formation of the genomic RNA dimer [320].

Both the TAR and polyA hairpin have been suggested to be involved in RNA packaging because mutation of these elements reduced the virion RNA content [116;128;171;317]. In these studies, the measured RNA content was normalized for the virion protein levels (RT and CA-p24). Subsequent experiments indicated that the amount of intracellular HIV-1 RNA is also reduced [317]. Therefore, the ratio of virion RNA to intracellular HIV-1 RNA seems a better measurement of the packaging efficiency than the ratio of virion RNA to virion protein. When the actual packaging efficiency was calculated, the 5' TAR hairpin was shown to moderately contribute to

packaging. For mutants with a destabilized 5' polyA hairpin, the reduced amount of virion RNA correlated perfectly with the reduction of intracellular HIV-1 RNA, which is caused by activation of the 5' polyA site [119]. Therefore, only the TAR hairpin remains a candidate accessory packaging signal.

FUTURE PROSPECTIVES

Although there is fairly good evidence for the secondary structure of RNA signals within the R region and the HIV-1 leader RNA in general, it should be emphasized that very little is known about the actual three-dimensional folding of these signals. It is realized that dealing with individual hairpins may be a gross over-simplification, because there may be structural or functional interactions between the different RNA modules [303]. The possibility of coaxial stacking between the 5' TAR and 5' polyA hairpin has already been mentioned. The RNA stretches between the stem-loop structures may also form long-distance interactions that contract the molecule into a more rigid structure. It is likely that some RNA domains may maintain a level of plasticity by being in equilibrium between two structures, and such RNA conformational transitions can provide unique regulatory possibilities.
Perhaps even a kind of quaternary RNA structure exists for HIV-1 RNA. It is generally believed that a mechanism exists whereby the 5' ends and 3' ends of the two genomes present in virions are held together in a conformation that allows efficient strand transfer. This higher order structure of the two viral RNA strands may facilitate strand transfer. Consistent with this notion is the finding that *in vitro* strand transfer with purified RT protein and nucleic acid factors is a rather inefficient process [41]. Furthermore, it was demonstrated that melting of the RNA dimer precluded strand transfer, but not strong stop DNA synthesis [245]. These results suggest that the conformation of the dimeric RNA genome facilitates the first strand transfer reaction of reverse transcription.

Elucidation of tertiary/ quaternary RNA structure and its role in regulated viral replication will not only present us novel molecular mechanisms of viral gene expression, but may also pave the way for recognition of similar mechanisms in the host cell. The detailed molecular understanding of retroviral replication may also

provide a basis for the development of novel antivirals. All these possibilities remain to be worked out in a variety of experimental systems, but it seems of utmost importance that structural information on the higher order structure of the HIV-1 RNA genome becomes available through biophysical studies.