J. Enzyme Inhibition, 1992, Vol. 6, pp. 9–34
Reprints available directly from the publisher
Photocopying permitted by license only

REVERSE TRANSCRIPTASE — A GENERAL DISCUSSION

ANTHONY L. DEVICO and M.G. SARNGADHARAN‡

Department of Cell Biology, Advanced BioScience Laboratories, Inc., 5510 Nicholson Lane, Kensington, Maryland 20895, U.S.A.

(*Received 2 October 1991*)

KEY WORDS: Reverse transcriptase, HIV, retroviruses, inhibitors.

1. INTRODUCTION

Retroviruses have a rather long history[1] but their recognition as a major human pathogen has been a slow development. Isolation of the first human retrovirus[2] and the identification of an etiological role for retroviruses in human diseases[3–8] have only been recent developments, but these signalled an unparalleled explosion of research activities in the 1980's that still continues into the new decade. The one major event that opened up the field of retrovirology was the discovery of reverse transcriptase (RT) as a key component of all retroviruses.[9,10] For the first time it provided a mechanism for the synthesis of a DNA copy of the viral genomic RNA. Integration of this DNA copy into cellular DNA is what establishes a retroviral infection. Since there is no mechanism available for selective and independent replication of this integrated DNA, called a *provirus*, or for its removal from cell DNA, a retroviral infection is considered to be a permanent event and during cell division the provirus is passed on to progeny as the chromosomal complement of DNA.

The life cycle of a retrovirus can be divided into two parts: first, the events that begin with the attachment of the virion on the cell surface and lead up to the integration of the provirus, and second, the steps that begin with the transcription of the provirus and end with the assembly and release of new mature virions from the cell surface. RT present in the confines of the incoming viral core is the key component that orchestrates the first part. While the molecular mechanisms that control the expression of the integrated provirus make use of normal cellular elements, the final phase of virus maturation does depend on viral gene products that are assembled at the cell surface. In the following section we will discuss the biosynthesis of RT, its role in virus replication and the molecular features contributing to some of its well recognized biochemical characteristics.

2. SYNTHESIS OF REVERSE TRANSCRIPTASE

2.1. Expression of Viral Genes Encoding RT

Reverse transcriptase is encoded by an open reading frame, designated *pol*, located in the center of the retroviral genome. The *pol* gene is preceded by the *gag* gene, which

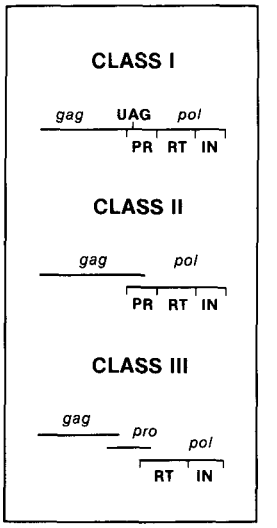

FIGURE 1 The genetic arrangement of *gag* and *pol* genes of various retroviruses. Class I: *gag* and *pol* genes in the same translational reading frame separated by a single termination codon. Class II: *pol* overlapping *gag* in the −1 reading frame. Class III: *gag* and *pol* separated by a third gene (*pro*) encoding the viral protease. Representative members of each class are given in the text.

encodes viral structural proteins, and is followed by the *env* gene, which encodes the viral surface glycoproteins. In addition to RT, the *pol* gene also encodes two other enzymes required for replication, the viral endonuclease/integrase (IN)[11–15] and in most cases the viral protease (PR).[16–19] Both mammalian and avian retroviral genomes encode these enzymes in the sequence PR–RT–IN.[14,15,17–20]

RT is expressed through the translation of a full length mRNA.[21] It is never synthesized as an individual protein, but is instead translated as part of a large polyprotein precursor[13,15,18,22–24] that contains the sequences of *gag* and *pol* gene products. The precursor, called gag-pol, is cleaved post-translationally to generate RT, along with the PR and IN enzymes.[16,24]

Despite the fact that gag and pol proteins are derived from a common precursor protein the retroviral genes are never arranged in the mRNA in a manner that allows uninterrupted translation of the precursor mRNA. A number of experiments have revealed that the expression of retroviral *pol* genes in fact depends upon several different mechanisms of translational suppression. Retroviruses can be divided into different classes based on which mechanism of suppression is used (Figure 1). The murine and feline leukemia viruses represent Class I, in which both the *gag* and *pol* genes are situated in the same translational reading frame in the mRNA, but are separated by a single amber UAG termination codon.[25,26] As a result, translation most frequently stops at this codon to produce the gag precursor polypeptide. Between 2% and 5% of the time, however, a translational read through event occurs and the gag-pol precursor is synthesized.[25–28] With the murine leukemia viruses, a glutamine residue is inserted at the position of the amber codon[25,27] and thus becomes the fifth residue in the amino acid sequence of the protease.[25]

This type of translational suppression has also been demonstrated with various model constructs *in vitro*.[28,29] These studies have revealed that suppression does not require any other *trans*-acting viral proteins, but seems instead to involve the recognition of intrinsic *cis*-acting components within the 300 nucleotides of viral mRNA adjacent to the codon.[29] Similar sequences derived from Moloney murine leukemia virus (Mo-MLV) effectively modulate suppression of UAA and UGA stop codons.[28] Therefore suppression is not dependent upon specific cellular suppressor tRNAs. Portions of the relevant downstream sequences are capable of forming stem-loop structures;[29] it is uncertain if such secondary structures influence suppression.

The avian retroviruses and the human immunodeficiency virus type 1 (HIV-1) represent Class II retroviruses, in which the 5′ end of the *pol* gene reading frame overlaps the 3′ end of the *gag* gene in the -1 direction.[30,31] In this arrangement, translation of *gag* gene sequences typically stop at a termination codon within the overlap region. Infrequently (between 5% and 10% of the time) the ribosome approaches the termination codon, slips back in the -1 direction [to a leucine codon in Rous sarcoma virus (RSV) and HIV-1[32,33]], and continues translation in the *pol* gene frame. As with the Class I mechanism of suppression, the process does not appear to require other viral proteins. Frame shift sites appear to be defined by specific heptanucleotide consensus sequences[34] and a specific secondary structure in the downstream RNA. This structure may be similar to the RNA pseudoknot shown to be necessary for frame shifting in the coronavirus system.[35] Alternatively, down stream stem-loop structures, which are theoretically possible, may influence the efficiency of frame shifting.[34,36] Such structures, however, do not seem to be required by HIV-1.[37]

The human T-lymphotropic viruses (HTLV-I and -II), bovine leukemia virus (BLV) and mouse mammary tumor virus (MMTV) represent Class III retroviruses, in which the protease is encoded by a separate gene that is positioned out of frame with respect to both *gag* and *pol*.[36,38-42] The expression of the *pol* gene therefore requires a -1 translational frame shift to move from the *gag* reading frame into the reading frame encoding the protease, and a second -1 frame shift to move from that reading frame into the *pol* frame. Although the frequency of each of the frame shifting events is as high as 25%,[36,40] the fact that two such events must occur in succession in order to express the gag-pro-pol protein results in a ratio of pol to gag proteins in Class III retroviruses that is lower than that observed for other retroviruses. While the translational control mechanisms of Mo-MLV and HIV-1 yield gag to gag-pol ratios of 20 : 1[25,27] and 8 : 1,[33] respectively, the gag to gag-pro-pol ratio in MMTV-infected cells is 30 : 1,[43] and is as high as 100 : 1 in the HTLV-II system.[44] These ratios may reflect what is necessary for maximal virion production.[45] It is possible that, due to spacial constraints within the viral core, a specific ratio of gag to pol proteins must be maintained during virus assembly. Thus the different suppression efficiencies observed among retroviruses may reflect differences in core structures. It is therefore possible that if compounds capable of interfering with suppression events could be developed, they may prove to be promising anti-retroviral agents. Such agents should be highly virus-specific since there are no known cellular genes that require termination suppression.

2.2 Processing of gag-pol Polyproteins

After translation gag-pol polyproteins are transported to the cell surface where, at

some point during virion assembly, they are proteolytically processed into the mature viral proteins.[18,24] Such cleavages are probably catalyzed entirely by the viral protease.[16,17,19,27] There is considerable evidence that processing occurs during and/or after budding of the virus.[17,19,46–48] This implies that the protease is capable of excising itself from the polyprotein precursor by an as yet undetermined autocatalytic mechanism. Mutant virions lacking protease function are capable of assembly and release, but contain uncleaved gag and gag-pol polyproteins.[17,48,49] The processing of polyprotein precursors is therefore not an obligatory step in the production of viral particles.

Precursor processing is apparently not required for the acquisition of RT-associated DNA polymerase activity. In studies on protease mutants of HIV-1[49,50] and Mo-MLV[17] unprocessed precursor proteins display a level of polymerase activity between 10% and 50% of that detected for the mature enzyme when measured using homopolymeric template primers. Although it cannot be ruled out that deletion mutations in the Mo-MLV protease domain may aberrantly activate the RT domain in the precursor, HIV mutants with single point mutations in the protease region also display substantial polymerase activity.[49,50] It therefore seems likely that wild type gag-pol precursors possess some DNA polymerase activity. It is significant that the Mo-MLV precursor can synthesize (-) strand DNA[17] in detergent permeabilized virions. This indicates that the precursor can not only bind its cognate tRNA correctly, but utilize it as a primer for DNA synthesis. This presents the intriguing possibility that the initial steps in replication may occur prior to the release of nascent virions.

Among various retroviruses RT is generated from the gag-pol precursor in a number of different ways. The processing pathways for three retroviruses are shown in Figure 2. The most simple processing occurs in the Mo-MLV system, where the precursor is fully cleaved to yield a monomeric 80 kD RT, IN and PR.[13,18,23,24,47,51] A noncovalent association between the Mo-MLV RT and IN may be maintained after such processing.[13]

In avian retroviral systems processing is more complex. In this case the *pol* encoded portion of the precursor includes only RT and IN since PR is encoded by the *gag* gene.[30,52] The pol segment is fully cleaved from the precursor to form a single polypeptide, designated β, which contains both RT and IN domains.[53–57] The β proteins form homodimers in which one of the proteins is subsequently cleaved[14,58] to form the smaller α subunit (representing the N-terminal portion of β) and a 32 kD phosphorylated IN protein, pp32 (from the β C-terminal portion), which possesses a specific endonuclease activity.[59,60] The mature avian myeloblastosis virus (AMV) RT is thus an $\alpha\beta$ heterodimer in which both proteins share common N-terminal sequences. Only the $\alpha\beta$ heterodimer displays full DNA polymerase and RNase H activities.[61]

The RT of the HIV-1 lentivirus is generated by yet another processing pathway. Initially, a 66 kD protein (p66), representing the RT domain, is completely cleaved from the IN and PR domains in the precursor.[62–64] Half of the p66 molecules subsequently undergo additional processing, apparently mediated by the viral PR,[62,63] which generates a 51 kD protein (p51) from the N-terminal portion of p66. It is uncertain whether processing involves p66 monomers or one of the proteins in a preformed p66 homodimer. In either case, the mature form of the enzyme found in virions is a heterodimer (p66/p51) in which both proteins share identical N-terminal sequences.[65–67] In this respect HIV-1 RT resembles the avian enzyme, although it should be emphasized that the former does not contain the IN domain in either of its subunits.

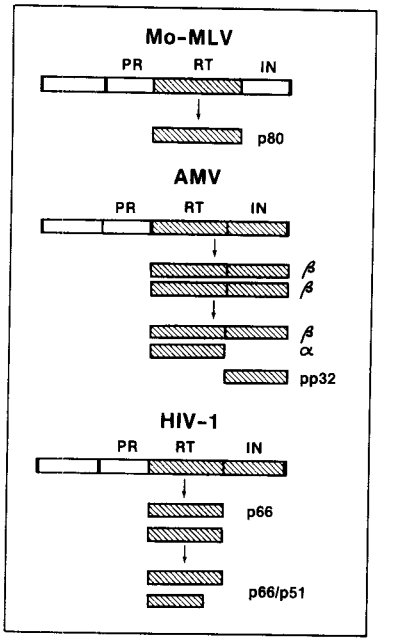

FIGURE 2 A schematic representation of gag-pol polyproteolytic processing yielding mature RT. Mo-MLV represents the process in murine retroviruses; AMV in avian retroviruses; HIV-1 in lentiviruses. Thick vertical lines mark the boundaries of the gag and pol regions; thin vertical lines divide the PR, RT and IN domains.

The p66 protein in itself displays high levels of both DNA polymerase and RNase H activities.[66–70] Mutagenesis studies on recombinant HIV-1 RT have demonstrated that most of the RNase H domain[64,71–73] and sequences important for polymerase activity[68–70,74] are located within the C-terminal fragment of p66. Thus p51 lacks RNase H activity and possesses only low levels of polymerase activity.[64,68–70,75] Since the cleavage of p66 inactivates the protein the purpose of such processing is not readily apparent. The relevance of the C-terminal p66 fragment is also obscure, since it does not exhibit RNase H activity in the absence of p51.[72,73] The p66/p51 heterodimer displays higher levels of enzymatic activity than p66, therefore processing may optimize folding and/or stability of the active site. The processing of HIV-2 RT is apparently very similar to that of HIV-1 RT. The purified enzyme is comprised of an equimolar mixture of 68 kD and 55 kD proteins.[76] An antibody directed against a sequence predicted to be in the C terminus of the enzyme reacts with only the p68 component, thus the p55 component is most likely a product of C-terminal cleavage of p68. An immunoaffinity purification procedure using the antibody recovers both proteins from virus extracts, indicating that the two components of HIV-2 RT are tightly associated. Like HIV-1 RT, polymerase activity is principally associated with the p68 component. It is notable that RTs from other lentiviruses, among them the equine infection anemia virus (EIAV),[77] and feline immunodeficiency virus (FIV)[78] also appear to undergo similar processing and heterodimer formation.

3. REPLICATION

Early in the life cycle of retroviruses, the viral RNA is reverse transcribed into a double-stranded DNA through an RNA · DNA hybrid intermediate. Over the past decade this process has been studied in great detail *in vitro* using virions that have been permeabilized by treatment with detergent under specialized conditions.[79–86] From these studies it is clear that the known enzymatic functions of RT, RNA-directed DNA polymerase, DNA-directed DNA polymerase, and ribonuclease H (RNase H), are in themselves sufficient to catalyze every reaction in the reverse transcription process. Most of the characteristics of these reactions have been studied in the murine and avian systems, and there is no compelling reason to believe that any other retrovirus replicates in a significantly different manner (Figure 3).

The retroviral virion contains a linked dimer of two 35 S (+) strand viral RNA molecules and about 70 RT molecules.[87–90] The RNA is blocked and methylated at the 5′ terminus[90] and polyadenylated at the 3′ terminus.[88] Attached to each RNA, near the 5′ terminus, is a tRNA molecule that serves as the initiator for synthesis of the (−) strand DNA[90–95] (Figure 3, Step A). The tRNA primer is bound, via a 16–18 nucleotide sequence in the 3′ acceptor stem, to a complementary sequence in the viral RNA called the primer binding site (pbs).[91,96–100] Depending on the retrovirus, only one species of tRNA is used as a primer. Thus RSV employs a $tRNA^{trp}$,[92,93] HIV-1 a $tRNA^{lys\,96}$ and Mo-MLV a $tRNA^{pro}$.[94,95]

The earliest product of reverse transcription is a discrete DNA species that is complementary to the viral RNA and is covalently attached to the tRNA[89,99–101] (represented by the thick line in Figure 3, Step B). This molecule, called (−) strong-stop DNA, is a product of the elongation of the tRNA primer up to the 5′ end of the genome.[79,91,99] Strong-stop DNA is a prominent species in the *in vitro* reaction, and therefore represents a point at which there is a pause in DNA synthesis by RT. Once the strong-stop DNA is made, it must switch templates and "jump" to a new position at the 3′ end of the genome in order for synthesis to continue.[102–105] Template switching is facilitated by a homologous repeat sequence (R), present in both ends of the genome.[106,107] The process also requires an RNase H-catalyzed degradation of the 5′ RNA template (represented by the dotted line in Figure 3, Step C) in order to expose the strong-stop DNA sequence complementary to R.[104,108–110] The R sequence in the 3′ end of the RNA then directs appropriate base pairing with the jumping DNA (Figure 3, Step D). The removal of the 5′ template appears to be critical to this process; Mo-MLV RT mutants that lack RNase H function cannot synthesize full-length proviral DNA.[111]

It is not known if RT directly promotes template switching via properties other than RNase H. Such an active role for HIV-1 RT is implied, however, by the fact that the enzyme can remain bound to DNA during switching between synthetic RNA templates.[112] RT may therefore help guide the strong-stop DNA from one template position to another. Nevertheless, template switching *in vivo* probably involves other viral proteins beside RT. This is indicated by the fact that the synthesis of full length (−) DNA in the *in vitro* endogenous reaction is highly sensitive to disruption of the virion core. Further, the formation of hairpin loops, mediated by an inverted repeat in the 5′ RNA terminus, interferes with template switching in reconstructed reactions containing RT in the absence of core proteins.[104] Hairpin structures do not form in reactions carried out with partially disrupted virions.

Since the viral genome contains two (+) RNAs, a persistent question has been

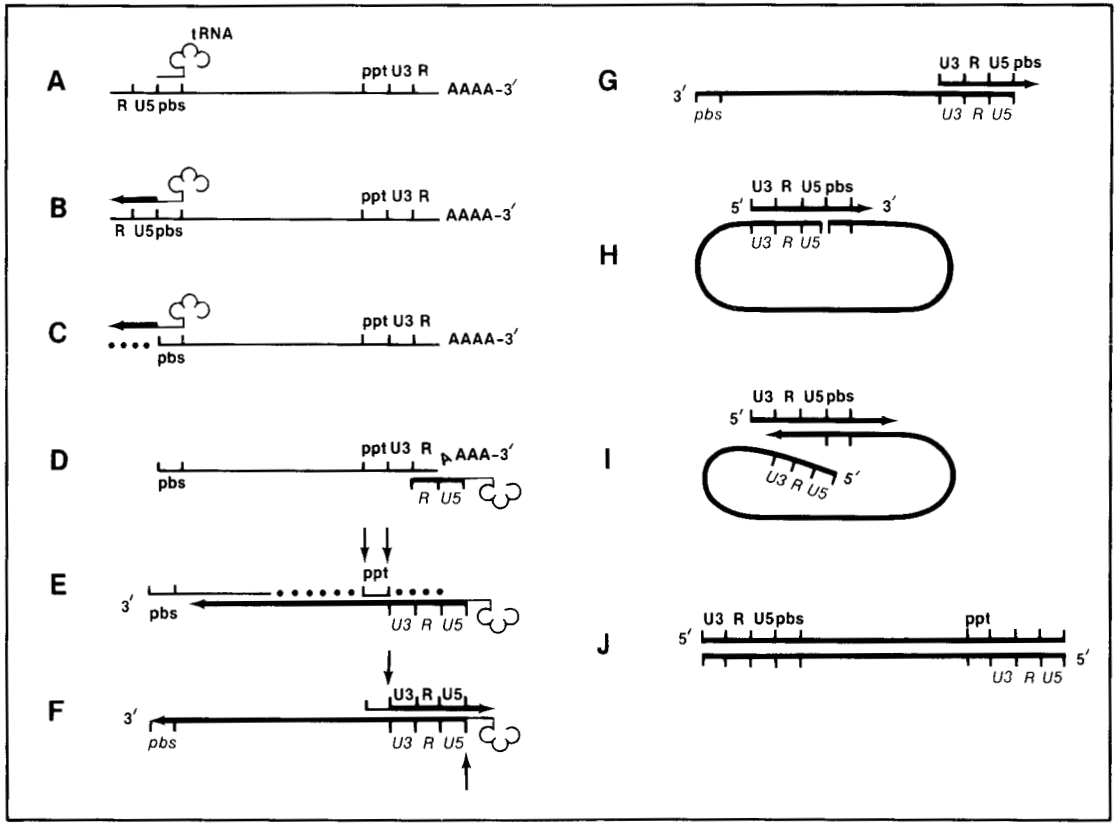

FIGURE 3 A schematic representation of the retroviral replication process. Thin lines represent RNA, thick lines DNA. Verticle arrows mark RNase H cleavage sites; dotted lines represent RNA that has been digested by RNase H. A detailed description of the process is given in the text. Genome regions identified in italics represent the minus strand.

whether template switching is an intrastrand (involving one RNA strand) or interstrand (involving both RNAs) event. This matter has been investigated recently using spleen necrosis virus particles containing two RNA species differing in restriction sites at their termini. These studies have revealed that the (−) strong-stop DNA template switch is an interstrand event.[113] How or why such an event is favored remains obscure. It may serve to promote genomic diversity, since the result is a "pseudo-diploid" provirus that contains genetic elements from both RNA strands.

After template switching has occurred, reverse transcriptase rapidly elongates the (−) strand DNA (Figure 3, Step E). The synthesis proceeds to the 5′ end of the template, which is now the pbs[79] because the rest of the genome was removed by RNase H prior to template switching. Under optimum conditions, RT can generate this full length DNA in 30 to 120 minutes.[114,115] It is likely that (−) strand synthesis is interrupted as reverse transcriptase encounters breaks in the RNA genome. The propensity of RT to remain bound to DNA may allow the enzyme to switch from the broken strand and continue synthesis on the other RNA template. Such template switching during a single round of replication would represent what has been termed a copy choice recombination event.[116–118]

Synthesis of the (−) strand DNA generates an RNA · DNA hybrid intermediate which becomes a substrate for RNase H.[110,119] Soon after the initiation of this synthesis RNase H removes the intact poly(A) tail by what is probably an endonucleolytic cleavage.[120] There is considerable evidence indicating that after this initial cut, RNase H continues to act on the nascent hybrid as an exonuclease (3′-5′) during subsequent (−) DNA synthesis (see Section 4).

Degradation of the hybridized RNA in essence creates both a free template and a primer for (+) strand DNA synthesis.[79,121] The primer is generated by selective cleavage of the RNA within a purine-rich sequence adjacent to the U3 region called the polypurine tract (ppt)[79,121–125] (indicated by the vertical arrows in Figure 3, Step E). The generation of the primer by RNase H can be readily duplicated *in vitro* using purified RT and model substrates.[125–130] The process appears to involve specific, sequence directed cleavages (see Section 4 for details).[125–129]

Elongation of the (+) strand DNA from the primer is carried out by the DNA-directed DNA polymerase activity of RT. This synthesis in fact begins even before the (−) strand DNA reaches full length.[114] This indicates that RNase H can generate the primer as soon as the relevant RNA · DNA hybrid is formed (such a process is depicted in Figure 3, Step E). Primer removal and primer extension take place in a single reaction as the (+) DNA is synthesized.[129] The primer is cleaved at a specific site by RNase H and is efficiently removed as a single fragment[126,129] (indicated by the upper arrow in Figure 3, Step F). The latter process generates the 5′ end of the (+) DNA U3 sequence. In some cases it is possible to detect initiation of (+) DNA strands at other points along the genome, both *in vivo* and *in vitro*.[124,131,132] Such initiations are in fact found at upstream sequences homologous to the polypurine tract.[133] The use of these incorrect initiation sites results in the generation of proviral DNA with a single stranded gap and abnormal termini.[132] Such molecules do not become integrated into the host cell DNA and therefore represent dead-end replication. Nevertheless, analyses of these molecules have further demonstrated that polypurine tract sequences are essential for the generation of (+) strand DNA primers by RNase H (see Section 4).

The (+) strand DNA is elongated toward the 5′ end of the (−) DNA strand, which is still attached to the tRNA. Synthesis terminates abruptly after the first 18 bases of

the tRNA primer are copied.[79,134] The stop probably corresponds to the point where RT reaches a modified tRNA base (an m^1A).[134] Transcription of the 18 tRNA bases yields a direct copy of the pbs which complements the sequence at the 3′ end of the (−) strand.[79] The species of (+) DNA generated by this synthesis is a prominent product in both *in vitro* and *in vivo* reactions,[103,135] and is called (+) strong-stop DNA (shown as the short thick line in Figure 3, Step F). In order for this molecule to be fully elongated, it must switch templates and jump to the 3′ end of the (−) strand. This occurs after the tRNA has been removed by RNase H at the RNA-DNA junction by what is probably an endonucleolytic cleavage[135] (indicated by the lower arrow in Figure 3, Step F).

After the tRNA has been removed (Figure 3, Step G), the exposed 3′ sequence in the (+) strong-stop DNA is free to switch templates and re-anneal to the complementary pbs in the 3′ end of the (−) strand. Unlike the first template switch, however, this transfer is an intrastrand event.[113] This pairing is thought to lead to the formation of a circular molecule[79,136] (Figure 3, Step H). Completion of the full length double-stranded DNA is accomplished by DNA-directed synthesis of each strand to the end of its template.[136,137] This requires another aspect of RT polymerase activity referred to as strand displacement synthesis (Figure 3, Step I). The final product of reverse transcription is a linear, double-stranded DNA (Figure 3, Step J) which, due to the two template switching events, is longer than the genomic DNA. This molecule is subsequently transported to the nucleus, where the IN protein mediates its integration.

4. RIBONUCLEASE H

4.1 General Characteristics of Ribonuclease H Activity

All known RT enzymes are associated with a ribonuclease H activity that, by definition, degrades only the RNA in RNA · DNA hybrid molecules.[99,138] This function is critical to several steps in the reverse transcription of the retroviral genome[139,140] (see Section 3 for details).

The RNase H activities of numerous retroviruses have been extensively characterized using a variety of synthetic substrates.[51,140–147] In every case the degradation of RNA does not depend on concurrent DNA synthesis since preformed hybrids serve as satisfactory substrates.[51,141–148] The products of the reaction using [^{3}H]poly(A) · poly(dT) as substrate are a series of oligonucleotides ranging from 1 to 20 residues.[51,138,142,147,148] Cleavages occur at the 3′ end of the 3′-5′ phosphodiester bond to yield 5′ phosphate and 3′ hydroxyl ends capable of serving as primer sites for DNA synthesis.[51,142,143,146,147]

Whether RNase H exhibits an endonucleolytic or exonucleolytic mode of action has been the object of some controversy. Early studies indicated that RNase H could not digest hybrid RNA blocked at both ends within closed circular molecules,[51,143,144,146] but could hydrolyze RNA under conditions where at least one end remained free. From these results it was concluded that retroviral RNase H acts as a processive exonuclease capable of attacking RNA from either the 3′ or 5′ end.[138,144,146]

Such a conclusion was at odds, however, with the complexity of RNase H substrates (and products) observed in virus-infected cells and in endogenous reactions carried out *in vitro*. Such studies clearly show that large segments of the viral RNA, such as the poly(A) tail, are released intact. It is particularly difficult to explain how

an exo-ribonuclease activity could remove intact the tRNA primer after (−) DNA synthesis. In this context the tRNA molecule assumes a complicated structure that is covalently linked to DNA at its 3′ end and is further hybridized to DNA. An endonucleolytic mode of cleavage clearly provides a more satisfactory explanation for several replication processes.

The issue has recently been resolved by a number of studies investigating the RNase H mediated cleavage of a variety of carefully designed synthetic substrates.[109,149–151] It has now been demonstrated that the RNase H activity associated with both HIV-1 and AMV RTs cleaves globin mRNA when the 5′ end is capped and the 3′ end is blocked by covalent modification.[150] Moreover, these enzymes can also cleave the RNA in relaxed, covalently closed RNA · DNA hybrid plasmids. Such results provide direct evidence that retroviral RNase H can mediate endonucleolytic cleavage of RNA lacking free ends and is therefore not strictly an exonuclease. The failure of early studies to demonstrate such a mode of action is likely a function of the supercoiled Col E1 hybrid plasmids and small ligated circles that were used as substrates.[51,143,144,146] In retrospect, these would seem to be particularly poor substrates for RNase H, since they contain only very short regions of hybridization (see below).

AMV RNase H has been investigated in considerable detail using a defined system in which a DNA primer is annealed to the internal portion of an end-labeled RNA template.[149] Cleavage of the hybridized RNA in the absence of DNA synthesis generates a 3′ fragment which remains invariant in size as the reaction proceeds, and a 5′ fragment which becomes progressively shorter. Nearly identical results have been obtained with HIV-1 and Mo-MLV RNase H when tested in similar systems.[150,151] Further analyses of the RNA species generated by these enzymes have demonstrated that the initial cleavage of internally hybridized RNA is achieved by an endonucleolytic RNase H activity (Figure 4, A). The RNase H cleavage sites do not display a random distribution but instead reflect a certain sequence selectivity.[148,152] The pattern of selectivity varies among RTs such that each enzyme generates a characteristic population of fragments from any given RNA template. Why some cleavage sites are preferred is not yet clear. The characterization of such sites may prove useful in the development of RNase H inhibitors.

After the initial endonucleolytic cleavage the mode of RNase H action changes, and the RNA is further hydrolyzed by 3′→5′ exonuclease activity.[151] Hydrolysis of the RNA is incomplete, however, and comes to a stop at a defined number of base pairs from the end of the hybrid (Figure 4, B). This length varies among RTs; with HIV-1 RT cleavage proceeds up to the last 7 base pairs,[151] and with AMV RT the last 11.[149]

AMV RNase H activity has also been studied under conditions of active DNA synthesis. During primer extension DNA polymerase and RNase H activities act simultaneously such that DNA elongation and RNA hydrolysis occur at the same rate. As a result only a short hybrid of relatively fixed length (7–14 base pairs with AMV RT) is maintained between the 3′ end of the nascent DNA chain and the 3′ end of the remaining template (Figure 4, C). Based on these results it has been suggested that a single RT molecule can carry out both RNase H and DNA polymerase activities, and that the length of the hybrid observed during DNA synthesis represents the distance between the active centers. However, the relationship between the RNase H and polymerase activities of AMV, HIV-1, and Mo-MLV RTs has recently been analyzed in greater detail using a challenged template assay that allows the measurement of synthesis and cleavage during a single RT-template primer interaction.[152] These experiments have demonstrated that once an RT molecule begins synthesizing

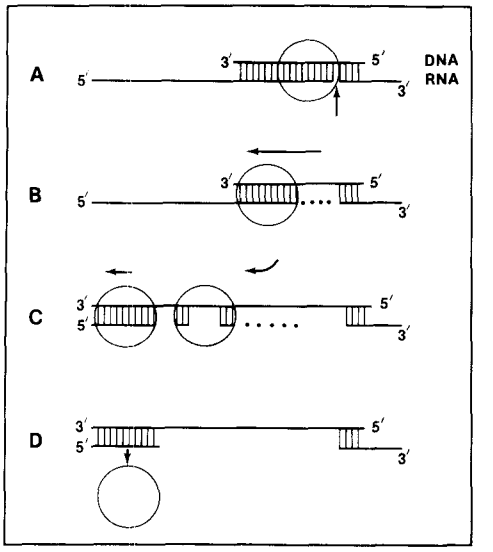

FIGURE 4 Cleavage of RNA · DNA hybrid molecules by RT-associated ribonuclease H. In the absence of DNA synthesis ribonuclease H first makes an endonucleolytic cleavage in the RNA strand of the hybrid (marked by the vertical arrow in A). Following the initial cleavage, the RNA is digested by a 3′ → 5′ exonuclease activity that stops at a fixed distance from the end of the hybrid (B). When deoxyribonucleotides are added to the system RNA digestion resumes as the DNA is elongated. Both processes occur simultaneously and are carried out by at least two RT molecules (C). Upon completion of DNA synthesis RT dissociates leaving a short fragment of RNA hybridized to the nascent DNA (D). Details of these processes are given in the text.

DNA, it is capable of making only random, infrequent RNA cleavages. Although the frequency of cleavage varies among RTs, in every case much of the potential substrate for RNase H is left undegraded. The polymerase and RNase H activities of RT are therefore not functionally coupled in a strict sense. Removal of the bulk of the hybrid RNA during DNA synthesis probably results from the RNase H activity of several enzyme molecules acting on the same template molecule simultaneously. The length of the hybrid observed during DNA synthesis may therefore represent the distance between the active centers in separate but coordinated RT molecules. (Such a situation is shown schematically in Figure 4, C).

After completion of DNA synthesis and dissociation of RT, a short undegraded fragment of RNA remains hybridized to the 3′ end of the nascent DNA (Figure 4, D). With AMV RT the length of this hybrid (7–14 base pairs) is the same length as the hybrid left after RNase H action in the absence of DNA synthesis. Since in either case the RNA is not digested, this length probably defines the minimum number of base pairs of an RNA · DNA hybrid required for RNase H activity. AMV RNase H activity is in fact not detected with 3 base pair hybrids until DNA synthesis extends the length of the hybrid to between 3 and 16 base pairs.[149]

In reconstructed reactions using Mo-MLV RT simultaneous DNA synthesis and RNA template removal have been demonstrated to promote template switching to an

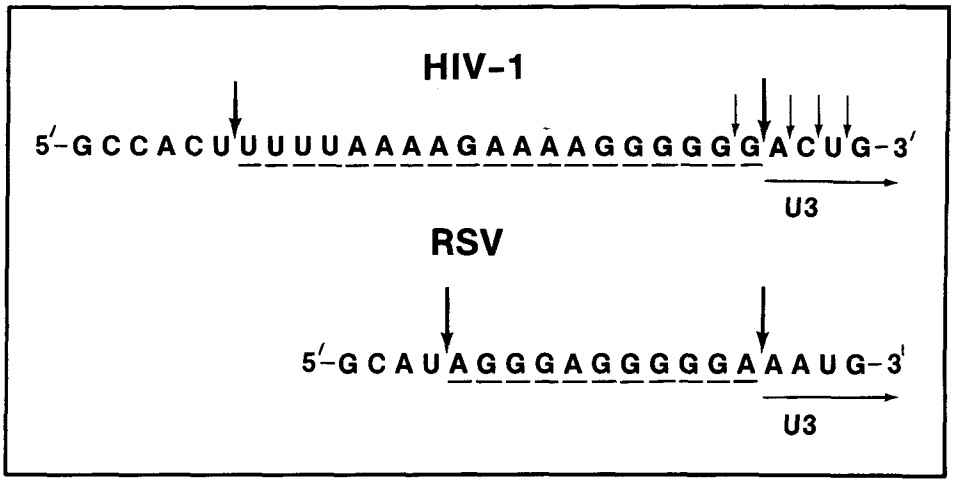

FIGURE 5 Generation of the (+) DNA primer by RNase H mediated cleavage. The polypurine tract sequences in HIV-1 and RSV viral RNAs are shown. Bold vertical arrows mark major RNase H cleavage sites; thin arrows mark minor cleavage sites made by HIV-1 RNase H. The primer sequence is marked with a dashed line. Horizontal arrow indicates direction of (+) DNA synthesis.

appropriate acceptor template.[109] It is therefore possible that a similar process promotes template switching after the (−) strong-stop DNA synthesis. Such a mechanism, however, predicts that a short fragment of viral RNA would remain bound to the 3′ end of the strong-stop DNA. How this would be removed is uncertain. Nevertheless, a capped, 12–15 base polynucleotide corresponding to the 5′ end of the viral RNA is a prevalent product of RNase H during (−) strong-stop DNA synthesis in the *in vitro* endogenous reaction.[108]

4.2. Generation of the (+) DNA Primer

It has been well established that during replication RNase H generates the RNA primer for (+) DNA synthesis. The primer is derived from a polypurine tract (ppt) in the viral RNA that is adjacent to the U3 region (see Section 3). The generation of the polypurine primer is notable because it involves considerable specificity of cleavage by RNase H. The process has been duplicated *in vitro* using a number of model substrates derived from relevant retroviral sequences.[127–130] Such systems have been extensively used to analyze the cleavage, extension and removal of the RNA primer. With purified HIV-1 RT all of these processes can take place in a single reaction, but are not strictly coupled, and can be carried out in individual reactions without loss of specificity.[129]

The polypurine primer is generated by specific cleavages within the 3′ and 5′ ends of the ppt (Figure 5). In the HIV-1 system[129] minor primer species of slightly different lengths are detected. These are generated by infrequent cleavages at secondary sites located within a few bases of the preferred site. The contribution of RNA sequence to the specificity of cleavage seems to be primarily at the level of sequence composition. A comparison of polypurine tracts from a diverse set of retroviruses reveals the conservation of a G at position -4 and a purine at position -1 relative to the (+) DNA

initiation site.[131] Specificity of cleavage also depends on the enzyme itself. The correct primer is not generated from the RSV ppt if Mo-MLV RT is used in the reaction.[128]

The generation of the polypurine primer by HIV-1 RT seems to involve endonucleolytic cleavage within the ppt.[130] It is possible that these cleavages occur in a coordinate fashion during (−) DNA synthesis but this has not yet been conclusively demonstrated. It is however clear that for some as yet unknown reason both the endo- and exonucleolytic functions of RNase H avoid, or are prevented from, cleaving within the polypurine primer sequence.

In the absence of DNA syntheses RNase H activity acts on RNA · DNA hybrids to generate a number of RNA fragments that may serve as potential primers. As a result the polypurine primer represents only a small fraction of the total.[128,129] Nevertheless, out of the population of RNA fragments extension of only the correct primer occurs. Surprisingly, this selectivity is independent of the polymerase used; correct (+) strand initiation can occur if Klenow polymerase is substituted for RSV RT.[123,128] The selection of the correct primer is therefore not a function of RT, and is probably due to a particularly stable interaction between the DNA and the RNase H generated polypurine primer.[128,153]

The removal of the primer from (+) DNA by AMV RT has been demonstrated to require that the RNA be hybridized to DNA.[154] Therefore the primer is released via an RNase H-mediated cleavage. This cleavage is almost certainly endonucleolytic, since HIV-1[129] and AMV[154] RTs release the primers intact. With HIV-1 RT the cleavage is very precise and occurs at a single site.[129] Due to the imprecision during primer formation, this cleavage is not necessarily at the RNA-DNA junction. As a result several ribonucleotides can be left on the 5′ end of the (+) DNA. In contrast, with AMV RT cleavage seems to occur precisely at the RNA-DNA junction.[154] Release of the primer by RT apparently involves specific sequence recognition. Mo-MLV RT inefficiently removes RNA primers from (+) DNA if synthesis is aberrantly initiated at upstream polypurine sites in the endogenous reactions.[131]

Even after generating the (+) DNA primer, RNase H activity probably remains essential to the completion of (+) DNA strand synthesis. The results of reconstituted reactions have demonstrated that HIV-1 RT cannot synthesize DNA directed by a DNA template in an RNA · DNA hybrid until the RNA has been removed. In this case synthesis and degradation do not appear to occur simultaneously,[112] RNA is first removed before synthesis proceeds.

The domain for RNase H appears to be located in the C terminal portion of most, if not all, RTs. Several lines of evidence support this conclusion. First, the C termini of numerous RTs share appreciable amino acid sequence homology with bacterial RNase H.[71] Second, deletion, linker insertion, and point mutations within the C termini of various recombinant RTs severely affect RNase H function.[70,139,155,156] Finally, the expressed C-terminal fragment of Mo-MLV RT in itself displays appreciable RNase H activity.[111]

The C-terminal portion of half of the HIV-1 RT p66 population is naturally cleaved from the protein during post-translational processing. At least one study reported that this fragment possessed a weak nonprocessive nuclease activity.[64] In contrast, a nearly identical polypeptide expressed in bacteria is only active when reconstituted with an N-terminal peptide of p66.[72] Nevertheless, the crystal structure of the C-terminal fragment shares a number of structural similarities with bacterial RNase H.[73] HIV-1 RT may represent a class of RTs, apart from Mo-MLV RT, in which certain N-terminal sequences are required in order to complete and/or stabilize the RNase H

domain. Similarly, sequences in the HIV-1 RNase H domain may play a reciprocal role in polymerase activity. An antibody raised against a sequence spanning amino acid residues 536–549 in the HIV-1 RT C terminus directly inhibits the template primer binding function of the enzyme.[74] The sequence is highly conserved among the RTs of lentiviruses, and thus the antibody is capable of directly inhibiting the polymerase activities of a number of these enzymes. The sequence recognized by the antibody is disordered in the crystal structure of the HIV-1 RNase H fragment, indicating that its correct positioning may be mediated by an interaction with one or both of the polymerase domains of the p66/p51 heterodimer.[73] The RNase H and polymerase domains of HIV-1 RT (and other lentiviral RTs) may therefore form a continuous groove in which RNA · DNA hybrid molecules are bound.

5. IMMUNOGENICITY OF REVERSE TRANSCRIPTASE

The reverse transcriptase enzymes of primate lentiviruses elicit a broad spectrum humoral immune response under conditions of natural infection. At least 80% of individuals infected with HIV-1 develop circulating antibodies that recognize both the p66 and p51 components of HIV-1 RT in immunoblot assays.[65,157] The RT proteins of the simian immunodeficiency virus (SIV) and HIV-2 display a similar level of immunogenicity.[76,158] A portion of the natural antibodies that bind these enzymes are capable of mediating direct inhibition of polymerase activity.[158–161] Such antibodies have been detected in 67% of randomly selected HIV-1-positive individuals.[158,160]

The significance of RT-inhibiting antibodies is obscure. The presence of these antibodies has been demonstrated to correlate with the inability to isolate HIV-1 from some individuals[160] and RT-inhibiting antibodies are reported to be most prevalent in individuals who are infected with HIV-1 but remain asymptomatic.[159] Such findings are notable since the overall level of antibody reactivity against RT does not vary significantly with clinical status.[157] The correlation between the presence of RT-inhibiting antibodies and the inability to isolate virus is intriguing, and suggests that the antibodies are capable of interfering with retroviral replication. How this might be achieved is obscure, however, since under natural conditions the RT is expected to be sequestered within the core of the virion and is therefore unavailable for binding. It is perhaps more likely that these antibodies are surrogate markers for other antibodies with reactivity against epitopes on the viral surface.[159]

Despite the fact that the HIV-1 and HIV-2 RTs are both highly immunogenic and share nearly 60% amino acid sequence homology, natural antibodies to these enzymes show little cross-reactivity.[76,158] Moreover, the majority of anti-HIV-2-positive sera have been found to be incapable of inhibiting HIV-1 RT.[158] Similarly, RT-inhibiting IgG purified from HIV-1-positive sera does not affect the activity of HIV-2 RT (Veronese and DeVico, unpublished observations). These results indicate that inhibiting antibodies are not directed against active site sequences, which are typically conserved. In support of this view, kinetic analyses have revealed that the mechanism of HIV-1 RT inhibition mediated by at least some of these antibodies is noncompetitive with respect to both substrate and template primer (Veronese and DeVico, unpublished observations). Therefore inhibitory antibodies interact with RT at sites other than those involved in substrate and template primer binding.

Monoclonal antibodies have been raised against HIV-1 RT that are capable of mediating direct inhibition of polymerase activity.[161] Such antibodies are similar to

their natural counterparts in that they do not cross-react with HIV-2 RT. Two of such monoclonals have been demonstrated to show a conformational requirement for their epitopes. It should now be possible to determine if natural RT-inhibiting antibodies compete with these monoclonals for binding with the same conformational epitope.

The high level of antibody response to RT observed in natural primate lentivirus infections is uncommon among other retroviruses. The only other known cases have been in AKR mice infected with Gross-MLV,[162] cows naturally infected with BLV,[163] and cats with feline leukemia virus.[164]

The mechanism through which anti-RT antibodies are elicited remains to be determined. There has been no compelling evidence for the presence of HIV-1 or HIV-2 RT on the surface of either virions or infected cells. Nevertheless, it is obvious that the immune system of some infected hosts are challenged by RT. The response is not necessarily limited to humoral immunity; a cytotoxic T-lymphocyte response against RT is elicited in some HIV-1-infected individuals.[165] Since the primate lentiviruses primarily infect lymphocytes, fragments of RT may be presented to the immune system in the appropriate context on the surfaces of these cells. An alternate explanation takes into account the fact that these viruses are highly cytopathic. Continuous lysis of infected cells might constantly challenge the immune system with either released RT or its gag-pol precursor.[157] Such a mechanism might further explain why the HIV-1 PR and IN, both *pol* gene products, are also highly immunogenic in infected individuals.[166,167]

6. DNA POLYMERASE

6.1. General Characteristics of DNA Polymerase Activity

The reverse transcriptase enzyme is capable of utilizing both RNA and DNA as a template to direct DNA synthesis.[51,168–175] This versatility sets RT apart from most other DNA polymerases and has made the enzyme the subject of intensive study. DNA synthesis by RT proceeds in the 5′ to 3′ direction,[168] requires a preformed template primer,[169] and is most efficient at physiologic salt, temperature and pH.[51,80,170–173] A particularly distinguishing feature of RT is its ability to synthesize DNA on poly (2′-*O*-methylcytidylate) template.[176,177] It should be stated at the outset that RT is not unique in its ability to catalyze RNA-directed DNA synthesis. One of the eukaryotic cellular enzymes, DNA polymerase γ, is also capable of utilizing RNA templates.[178] DNA polymerase γ, however, differs from RT in several respects, and cannot use poly (2′-*O*-methylcytidylate) as a template.[175–178]

In reconstructed reactions RT is able to efficiently transcribe both homopolymeric and heteropolymeric templates.[51] Synthesis is processive on RNA templates[112,173] and is initiated most efficiently when DNA oligomers are used as primers.[179] In addition to common properties, various RTs display certain distinguishing features depending on the template primer and divalent cation used. For example, the Mo-MLV RT is most active using $(dT)_{\sim 15} \cdot (A)_n$ template primers and Mn^{2+} as divalent cation.[51,80] In contrast, HTLV-I and HIV-1 RTs are most active on the $(dG)_{\sim 15} \cdot (C)_n$ template primer using Mg^{2+}.[171,174,175]

The mechanism of DNA synthesis carried out by HIV-1 RT has been examined in detail through kinetic and processivity analyses.[112,180–182] Several characteristics of the reaction are notable. The reaction is ordered, with free enzyme binding template primer first and deoxyribonucleoside triphosphates second.[180–182] Such an ordered

mechanism is not unusual, being followed by several other DNA polymerases. The K_{on} value for the binding of enzyme to $(dT)_{\sim 15} \cdot (A)_{810}$ ($\sim 10^9\,M^{-1}\,S^{-1}$) has been determined to be similar to the theoretical diffusion-controlled rate for molecules of such size.[180,182] Binding of template primer appears to occur initially through the primer moiety since it has been demonstrated that $(dT)_{15}$, as well as certain modified polydeoxyribonucleotides, act as competitive inhibitors.[180,181] Once formed, the RT-template primer complex displays a half-life of 100–200 seconds, and a dissociation constant of 2–3 μM.[112] Chain elongation during DNA synthesis proceeds at a rate of about 4–10 nucleotides per second.[112,181] Synthesis on RNA templates is processive, the longest primer extensions are typically greater than 300 nucleotides with an average processivity of about 100 nucleotides.[112,181] When $(dT)_{14} \cdot (A)_{810}$ is used as template primer, termination after incorporation of the first dTMP residue is observed to occur greater than 20 times more frequently than after subsequent additions.[181] This observation is consistent with a reaction mechanism in which the initiation of synthesis is kinetically distinct from subsequent elongation steps.

The K_{off} value for enzyme-template primer dissociation during the initial reaction step has been calculated to be 1000-fold greater than the K_{off} value during processive synthesis.[182] The mechanism for the K_{off} decrease as synthesis proceeds is not clear, but may involve additional and/or different interactions than are involved in initial template primer binding. Recent evidence indicates that the increase in processivity observed after the $(dT)_{14}$ primer is elongated by one residue may in fact be due to a change in the rate constant for the formation of the enzyme-template primer complex.[183] The estimated K_d values for $(dT)_n \cdot (A)_{1000}$ have been reported to be 20–30 nM when $n = 10$–14 and 0.11–0.14 nM when $n = 16$–20. Thus an abrupt change in K_d, as well as in the pattern of primer extension, is observed between primer lengths of 14 and 16 nucleotides. It is interesting that such a change is observed with the p66/p51 heterodimer and the p66 homodimer forms of the enzyme, but not with p51 homodimers. This indicates that the RNase H domain in the C-terminal portion of p66 may contribute to the binding of template primers in which the primer is greater than 14 nucleotides in length.

DNA synthesis by HIV-1 RT on $(dT)_{20} \cdot (dA)_{380}$ template primer is nearly distributive. Most primers are extended by only 1–20 nucleotides before termination occurs.[112] Processivity on DNA templates of random sequence is considerably greater, on average about 50 nucleotides. The strong stops on such templates do not occur at random but are located within 1 or 2 bases after at least three *A* or *T* residues. This correlates with the distributive nature of the RT on the polydeoxyadenylate template. Such results suggest that the mechanism for promoting increased processivity on RNA templates cannot operate efficiently on some DNA templates.

According to the favored model for retroviral replication (Figure 3), synthesis of the (+) DNA strand must proceed through a region of double stranded DNA. Since RT has no DNA exonuclease activity, synthesis requires the displacement of the non-template strand downstream from the polymerization site. Such a process, called strand displacement synthesis, has been directly demonstrated with HIV-1 RT in reactions containing double stranded DNA templates of random sequence.[112] A 4 base gap in one of the strands is sufficient to initiate elongation of the DNA strand upstream from the gap. The observed extensions are due to processive strand displacement synthesis, which proceeds for about 50 base pairs. Since no panhandle structures are detected among the DNA products, strand switching apparently does not occur during displacement synthesis. The displacement of RNA from DNA templates

during DNA synthesis has not been demonstrated for RT. Synthesis instead appears to rely entirely upon the removal of the RNA by RNase H activity (see Section 4).

6.2. Fidelity of DNA Synthesis by RT

It has long been established that RT enzymes are significantly more error-prone than many other DNA polymerases when copying both RNA and DNA templates.[184–187] These enzymes have further been demonstrated to lack a 3′ to 5′ exonuclease (proof-reading) function, and therefore cannot excise and correct misincorporated nucleotides.[188,189] The fidelity of synthesis varies among RTs, with HIV-1 RT demonstrating the lowest level of accuracy to date.[189–191]

The error specificities of several RTs have recently been determined at a large number of sites within $M13_{mp}2$ DNA templates.[192,193] The dominant errors are single base substitutions and single base frameshifts, and are clustered around specific mutational hot spots. The distribution pattern of such hot spots varies among the RTs examined.[192,193] HIV-1 RT introduces both +1 and −1 frameshifts within runs of a common base. These probably arise due to a mechanism of strand slippage between template and primer.[192] AMV RT introduces only −1 frameshifts which sometimes occur at nonrun sequences. In this case the misinserted base is thought to shift along the template and form a base pair at some other proximal site.[193,194] Once base pairing occurs, the error is introduced by extension of the DNA. Single base substitutions by HIV-1 RT have been demonstrated to arise due to the propensity of the enzyme to elongate mispaired nucleotides located at the 3′ termini of nascent DNA strands.[195] With HIV-1 RT most substitution errors are located at the boundaries of single base runs, therefore mispaired termini probably originate by a transient slippage mechanism similar to the one that produces single base frameshifts.

One point should be made concerning $M13_{mp}2$ DNA templates and error-prone replication. As mentioned earlier, DNA template-directed DNA synthesis is highly distributive and results in frequent termination during chain elongation. This fact appears to influence the error-rate, since one base frameshifts are found to occur more often at $M13_{mp}2$ template runs with the highest probability of termination.[192] DNA synthesis directed by RNA templates is more processive and may therefore display a somewhat different pattern and specificity of errors. It is possible that the specificity of errors introduced during retroviral replication may depend upon which DNA strand is being transcribed.

The reduced fidelity of DNA synthesis by RT, coupled with its inability to correct errors, is almost certainly relevant to the fact that retroviruses display a high rate of genomic mutation[196–198] and are among the fastest mutating viruses. The error rate of AMV RT measured in forward mutation assays is in fact sufficient to account for the AMV mutation rate observed during a single round of replication.[193,197] In other viral systems the mutation rate has been correlated with the genomic variability observed in nature.[199] Therefore the unusually high level of variability observed among natural isolates of HIV-1[200,201] can probably be explained in part by the fact that HIV-1 RT is 2 to 10 times less accurate than any other RT examined.[190,191] Such a mechanism for generating diversity may also function in other retroviral systems. Isolates of EIAV also display a high level of genomic heterogeneity similar to that observed with HIV-1. A preliminary analysis of the fidelity of DNA synthesis by EIAV RT indicates that it is as error prone as HIV-1 RT.[77]

7. INTERACTIONS WITH tRNA PRIMERS

7.1. Specificity of tRNA Binding

The virions of various retroviruses contain several unique species of cellular tRNA. A portion of these molecules become annealed to the viral RNA and serve as primers for viral DNA synthesis. The interactions between such tRNA primers and the RT molecule have been studied in considerable detail. Such studies have been greatly facilitated by the fact that stable RT-tRNA complexes can be readily generated *in vitro* using only purified enzyme and the appropriate tRNA.

Various RTs display a specific binding affinity for only the species of tRNA that serves as primer. AMV RT can form a selective and stable interaction with the avian $tRNA^{trp}$ primer species even in the presence of competing cellular tRNAs.[202-204] Similarly, HIV-1 RT can selectively bind its $tRNA^{lys,3}$ primer species from a mixture containing a 100-fold excess of $tRNA^{trp}$ and $tRNA^{pro}$.[205] Such selective binding, however, may not be a universal property of all RTs. Mo-MLV RT does not demonstrate a specific interaction with its primer $tRNA^{pro}$ under *in vitro* conditions, and instead binds a variety of tRNA species non-selectively.[203-206]

The binding of tRNA is highly dependent on the conformation of the molecule. The necessary conformation is apparently not altered by aminoacylation; AMV RT is capable of binding both charged and uncharged $tRNA^{trp}$.[204,207] Nearly the entire native conformation seems to be required for binding, since AMV RT fails to interact with even large 3′ or 5′ fragments of the $tRNA^{trp}$ molecule.[204,208] Only the eight 3′ terminal residues in the acceptor stem have been demonstrated to be dispensable for binding.[209] In contrast, the structural requirements for tRNA binding are less stringent once the molecule is base-paired with the RNA template. DNA synthesis by AMV RT can be initiated by a small, but specific, portion of $tRNA^{trp}$ corresponding to the 3′ acceptor stem and $T\psi C$ loop.[208] The relaxation of structural requirements for the binding of annealed tRNA probably explains why AMV RT can efficiently transcribe 70s RNA from non-avian retroviruses.[204] The recognition of annealed tRNA does require at least some structural features, however, since very short fragments of $tRNA^{trp}$ fail to serve as primers even though they are base paired to the template and provide a free 3′ end.[208]

The tRNA domain(s) recognized by RT has been partially mapped by photochemically cross-linking the molecules, subjecting the tRNA to limited nuclease digestion, and sequencing the RNA fragments that remain bound to RT. HIV-1 RT has been shown to cross-link with oligonucleotide fragments of $tRNA^{lys,3}$ derived from the 3′ portion of the anticodon stem and the anticodon loop.[205] Similarly, AMV RT components cross-link with fragments derived from the $tRNA^{trp}$ anticodon stem and loop and the 3′ end of the acceptor stem[210] (Figure 6, A). In the presence of enzyme, however, the anticodon loop remains sensitive to pancreatic ribonuclease A and nicking of the loop by S1 nuclease does not prevent interaction with RT.[211] Thus the enzyme does not appear to interact with the entire anticodon stem and loop structure. Direct interactions may occur at the anticodon stem, since AMV RT protects this portion of the tRNA from digestion by cobra venom nuclease and *Neurospora crassa* endonuclease. Other putative contact sites on $tRNA^{trp}$ have been mapped using various chemicals and nucleases. AMV RT protects the 3′ side of the DHU stem and the 5′ side of the $T\psi C$ stem from alkylation of ethylnitrosourea.[212] The

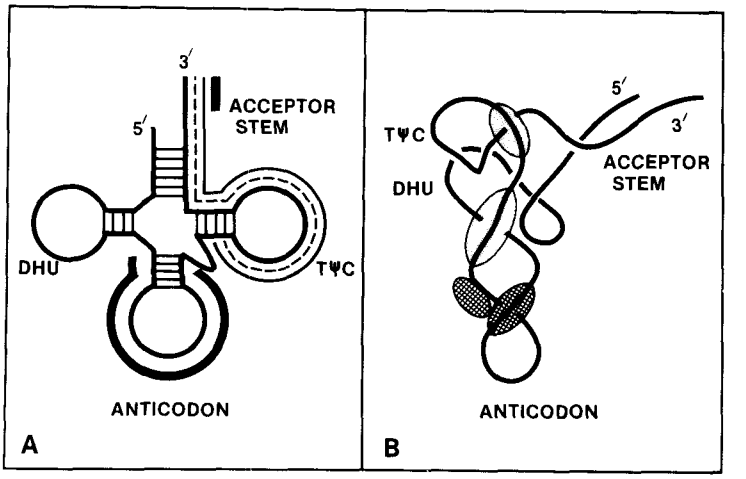

FIGURE 6 Reverse transcriptase binding domains within the tRNA primer molecule. A. tRNA depicted in the cloverleaf arrangements. The dashed line marks the portion of the molecule that anneals to the viral RNA primer binding site. The thin line corresponds to the minimal portion of tRNA required to initiate DNA synthesis. The thick line corresponds to the portions of tRNA that photochemically cross-link to RT. B. L-shaped structure of tRNA. Interactions with RT protects the anticodon stem from nuclease digestion (dark shading) and the DHU stem and TψC stem from alkylation (light shading).

TψC loop may also represent a contact site since cyanoethylation of the pseudouridine residues within the loop destroys the ability of $tRNA^{trp}$ to bind to the RT.[211] It is notable that all of these putative contact sites are located on the outer portion of the normal L-shaped structure of tRNA[212] (Figure 6, B). It has been proposed that much of this structure is likely to be maintained by tRNA after it has annealed to viral RNA.[213]

Since the tRNA binding exhibited by several RTs is selective for the priming species, interaction with the molecule probably depends upon primary sequence as well as structure. One or both of these characteristics can be shared by multiple tRNA species, which probably explains why AMV RT also binds yeast,[214] murine,[211] and bovine[207] $tRNA^{trp}$, and an avian $tRNA^{met}$ species,[202–204] with appreciable affinity. In contrast, the Mo-MLV enzyme seems to have lost the ability to recognize any specific tRNA sequences, and may instead recognize structural features conserved among all tRNAs. In any case, a more complete understanding of the tRNA sites recognized by RT could potentially be exploited for the development of replication inhibitors.

It is significant that RT enzymes bind the 3′ acceptor stem of tRNA, since this portion of the molecule anneals to the viral RNA template.[95,213] A portion of the *TψC* loop, which also appears to interact with RT, may also base pair with the template[105] (Figure 6, A). There is evidence that AMV RT melts, or unwinds, the tRNA structure in order to facilitate this base pairing at biologically relevant temperatures. In the presence of enzyme the 3′ acceptor stem becomes susceptible to RNAse T_1 cleavage and the residue at position 10 displays enhanced reactivity with ethylnitrosurea. This indicates the occurrence of a fundamental event in which the RNA duplex of the acceptor stem has become unwound.[212,215] Binding to RT also causes position 27 in the anticodon stem to become accessible to cleavage by cobra venom nuclease and

position 36 in the anticodon loop to undergo spontaneous degradation.[212] These are probably distal effects, since these portions of the molecule do not appear to be in direct contact with RT.[211,212] The enzyme therefore seems to destabilize regions of tRNA where most tertiary interactions maintaining structure are located.[212] Presumably because AMV RT can facilitate tRNA melting and annealing to the template, DNA synthesis can be demonstrated *in vitro* using only enzyme, $tRNA^{trp}$ and purified 35S RNA.[215,216]

Other RT enzymes appear to contrast with AMV RT in that they are apparently uninvolved in the process of base pairing tRNA primers to the viral RNA template. Although HIV-1 RT separately binds both viral RNA and tRNA, it cannot synthesize viral DNA in the absence of the nucleocapsid protein (p15).[205] Further, certain Mo-MLV mutant virions lacking functional RT nevertheless contain primer $tRNA^{pro}$ that is properly base-paired with the viral RNA.[217] By analogy to HIV-1, the Mo-MLV nucleocapsid protein may promote the annealing process instead of RT.

7.2. RT Structure and tRNA Binding

Both the α and β subunits of AMV RT become cross-linked to $tRNA^{trp}$ after photochemical treatment.[210] This would indicate that both subunits of the enzyme interact with, or are proximal to, portions of the tRNA. When AMV RT is reacted with periodate-treated $tRNA^{trp}$, a Schiff's base forms between a lysine in the α subunit and the derivitized 3′-terminal ribose of the tRNA. After longer reaction times the β subunit also becomes linked to the tRNA.[214] The extent of primer linkage is not reduced by nonprimer tRNAs and therefore reflects the specificity of AMV RT for $tRNA^{trp}$ The α subunit, however, does not in itself bind $tRNA^{trp}$ with any discernible affinity.[207,218] This correlates with an inability of the α subunit to efficiently transcribe 70S viral RNA.[204,218,219] In contrast, both the $\alpha\beta$ and β subunit forms of the enzyme are capable of establishing a strong interaction with $tRNA^{trp}$.[218] Taken together these results may be interpreted to indicate that the β subunit is responsible for the initial binding of tRNA, and that the α subunit later establishes secondary interactions. The β subunit may also act to stabilize such secondary interactions. Both subunits probably cooperate in binding a single tRNA molecule, since complexes contain a 1 : 1 molar ratio of tRNA to enzyme.[207]

In the case of HIV-1 RT, both the p66 and p51 subunits of the enzyme become cross-linked to the $tRNA^{lys,3}$ anticodon stem and loop.[205] However, neither subunit in purified form can form the specific complex with $tRNA^{lys,3}$. Both proteins probably act coordinately to form a single tRNA binding site since complexes contain a 1 : 1 molar ratio of $tRNA^{lys,3}$ to RT heterodimer.

7.3. Selective Encapsidation of Primer tRNA

Retroviral virions are typically enriched in their primer tRNAs even though such molecules make up only a small fraction of total cellular tRNA.[220–222] Since the number of RT and tRNA molecules in the virion are about equal,[220,223] it is probable that this selective encapsidation is a reflection of RT binding the primer tRNA in the infected cell and transporting the molecule into the virion. ASV mutants which lack functional RT are in fact non-selective in the incorporation of $tRNA^{trp}$ into virions.[224] In contrast, mutants containing normal RT but lacking a packaged viral genome selectively incorporate normal amounts of the primer tRNA. Mo-MLV mutants

lacking RT also exhibit a decrease in the level of viron-associated $tRNA^{pro}$.[217] Such results strongly indicate that Mo-MLV RT may in fact display some specificity for $tRNA^{pro}$ binding *in vivo*. Mo-MLV virions are enriched in $tRNA^{pro}$ only 3-fold relative to the proportion in cellular RNA, while in avian virions the primer $tRNA^{trp}$ is enriched more than 30-fold.[225] It is possible that the Mo-MLV RT-$tRNA^{pro}$ interaction may be relatively weak, and therefore difficult to duplicate *in vitro*.

8. CONCLUSION

The RT enzyme is clearly essential for the completion of the retroviral life cycle. This fact makes RT a primary target for compounds that could be used in treating diseases of retroviral etiology such as AIDS. The nucleoside analogues 3′-azido-thymidine (AZT) and dideoxyinosine (DDI), which inhibit the polymerase function of HIV-1 RT by causing chain termination, have already been used to successfully treat HIV-1 infection. Nucleoside analogues, however, also affect other DNA polymerases and are therefore inherently toxic to the cell. Fortunately, the continued study of RT has revealed a number of other features that could be targeted in order to control retroviral infection. The biosynthesis of RT, for example, could be targeted at the level of termination suppression (see Section 2.1) during translation of the viral mRNA. Agents capable of interfering with this process are expected to be particularly specific anti-retroviral compounds, since there are no known cellular processes that require termination suppression. Alternatively, compounds might be developed that exploit the fact that RT-associated RNase H activity preferentially acts at certain RNA sequences. Analogues derived from such sequences may prove to be effective RNase H inhibitors. The specificity of tRNA primer binding by RT provides yet another feature that could be targeted by anti-retroviral compounds. It is easy to imagine how agents that interfere with RT-tRNA primer interactions could effectively block replication. Such agents may take the form of tRNA primer molecules in which the 3′ end is blocked, or perhaps be comprised of key portions of the tRNA molecule whose competitive binding with RT could lead to a block in the selective enrichment and encapsidation of the specific primer tRNA species during virus assembly (see Section 7.3). In any case, the intense study of RT in recent years should now make it possible to rapidly develop and test new drugs that either act as RT inhibitors or block the biosynthesis and function of the RT molecule.

Acknowledgement

We wish to thank Joan Jackson and Sharon Orndorff for editorial assistance. This work is in partial fulfillment of the requirement for a Ph.D. degree for Anthony L. DeVico from George Washington University Department of Biology, Genetics Program.

References

1. Ellerman, V. and Bang, O. (1908) *Zentralbl. Bacteriol. Mikrobiol. Hyg.*, **46,** 595.
2. Poiesz, B.J., Ruscetti, F.W., Gazdar, A.F., Bunn, P.A., Minna, J.D. and Gallo, R.C. (1980) *Proc. Natl. Acad. Sci. USA*, **77,** 7415.
3. Kalyanaraman, V.S., Sarngadharan, M.G., Nakao, Y., Ito, Y., Aoki, T. and Gallo, R.C. (1982) *Proc. Natl. Acad. Sci. USA*, **79,** 1653.
4. Robert-Guroff, M., Nakao, U., Notake, K., Ito, Y., Sliski, A. and Gallo, R.C. (1982) *Science*, **215,** 975.

5. Barre-Sinoussi, F., Chermann, J.-C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauquet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W. and Montagnier, L. (1983) *Science*, **220,** 868.
6. Gallo, R.C., Salahuddin, S.Z., Popovic, M., Shearer, G.M., Kaplan, M., Haynes, B.F., Palker, T.J., Redfield, R., Oleske, J., Safai, B., White, G., Foster, P. and Markham, P.D. (1984) *Science*, **224,** 500.
7. Popovic, M., Sarngadharan, M.G., Read, E. and Gallo, R.C. (1984) *Science*, **224,** 497.
8. Sarngadharan, M.G., Popovic, M., Bruch, L., Schüpback, J. and Gallo, R.C. (1984) *Science*, **224,** 506.
9. Temin, H.M. and Mizutani, S. (1970) *Nature* (*Lond.*), **27,** 226.
10. Baltimore, D. (1970) *Nature* (*Lond.*), **27,** 226.
11. Donehower, L.A. and Varmus, H.E. (1984) *Proc. Natl. Acad. Sci. USA*, **81,** 6461.
12. Schwartzberg, P., Colicelli, J. and Goff, S.P. (1984) *Cell*, **37,** 1043.
13. Hu, S.C., Court, D.L., Zweig, M. and Levin, J.G. (1986) *J. Virol.*, **60,** 267.
14. Copeland, T.D., Grandgenett, D.P. and Oroszlan, S. (1980) *J. Virol.*, **36,** 115.
15. Eisenman, R.N., Mason, W.S. and Linial, M. (1980) *J. Virol.*, **36,** 62.
16. Yoshinaka, Y. and Luftig, R.B. (1977) *Cell*, **12,** 709.
17. Crawford, S. and Goff, S.P. (1985) *J. Virol.*, **53,** 899.
18. Copeland, T.D., Gerard, G.F., Hixson, C.W. and Oroszlan, S. (1985) *Virology*, **143,** 676.
19. Oroszlan, S. and Luftig, R.B. (1990) *Curr. Top. Microbiol. Immunol.*, **157,** 153.
20. Levin, J.G., Hu, S.C., Rein, A., Messer, L.I. and Gerwin, B.I. (1984) *J. Virol.*, **51,** 470.
21. Weiss, S.R., Varmus, H.E. and Bishop, J.M. (1977) *Cell*, **12,** 983.
22. Oppermann, H., Bishop, J.M., Varmus, H.E. and Levintow, L. (1977) *Cell*, **12,** 993.
23. Kopchick, J.J., Jamjoom, G.A., Watson, K.F. and Arlinghaus, R.B. (1978) *Proc. Natl. Acad. Sci. USA*, **75,** 2016.
24. Eisenman, R.N. and Vogt, V.M. (1978) *Biochim. Biophys. Acta*, **473,** 187.
25. Yoshinaka, Y., Katoh, I., Copeland, T.D. and Oroszlan, S. (1985) *Proc. Natl. Acad. Sci. USA*, **82,** 1618.
26. Yoshinaka, Y., Katoh, I., Copeland, T.D. and Oroszlan, S. (1985) *J. Virol.*, **55,** 870.
27. Yoshinaka, Y. and Luftig, R.B. (1980) *J. Gen. Virol.*, **48,** 329.
28. Feng, Y.X., Levin, J.G., Hatfield, D.L., Schaefer, T.S., Gorelick, R.J. and Rein, A. (1989) *J. Virol.*, **63,** 2870.
29. Schwartz, D.E., Tizard, R. and Gilbert, W. (1983) *Cell*, **32,** 853.
30. Ratner, L., Haseltine, W., Patarca, R., Livak, K.J., Starcich, B., Josephs, S.F., Doran, E.R., Rafalski, J.A., Whitehorn, E.A., Baumeister, K. and et al, (1985) *Nature* (*Lond.*), **313,** 277.
31. Jacks, T. and Varmus, H.E. (1985) *Science*, **230,** 1237.
32. Panganiban, A.T. (1988) *J. Virol.*, **62,** 3574.
33. Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J. and Varmus, H.E. (1988) *Nature* (*Lond.*), **331,** 280.
34. Jacks, T., Madhani, H.D., Masiarz, F.R. and Varmus, H.E. (1988) *Cell*, **55,** 447.
35. Brierley, I., Digard, P. and Inglis, S.C. (1989) *Cell*, **57,** 537.
36. Jacks, T., Townsley, K., Varmus, H.E. and Majors, J. (1987) *Proc. Natl. Acad. Sci. USA*, **84,** 4298.
37. Wilson, W., Braddock, M., Adams, S.E., Rathjen, P.D., Kingsman, S.M. and Kingsman, A.J. (1988) *Cell*, **55,** 1159.
38. Hiramatsu, K., Nishida, J., Naito, A. and Yoshikura, H. (1987) *J. Gen. Virol.*, **68,** 213.
39. Hizi, A., Henderson, L.E., Copeland, T.D., Sowder, R.C., Hixson, C.V. and Oroszlan, S. (1987) *Proc. Natl. Acad. Sci. USA*, **84,** 7041.
40. Moore, R., Dixon, M., Smith, R., Peters, G. and Dickson, C. (1987) *J. Virol.*, **61,** 480.
41. Rice, N.R., Stephens, R.M., Burny, A. and Gilden, R.V. (1985) *Virology*, **142,** 357.
42. Shimotohno, K., Takahashi, Y., Shimizu, N., Gojobori, T., Golde, D.W., Chen, I.S., Miwa, M. and Sugimura, T. (1985) *Proc. Natl. Acad. Sci. USA*, **82,** 3101.
43. Dickson, C. and Atterwill, M. (1979) *Cell*, **17,** 1003.
44. Mador, N., Panet, A. and Honigman, A. (1989) *J. Virol.*, **63,** 2400.
45. Felsenstein, K.M. and Goff, S.P. (1988) *J. Virol.*, **62,** 2179.
46. Witte, O.N. and Baltimore, D. (1978) *J. Virol.*, **26,** 750.
47. Tanese, N., Roth, M.J. and Goff, S.P. (1986) *J. Virol.*, **59,** 328.
48. Katoh, I., Yoshinaka, Y., Rein, A., Shibuya, M., Odaka, T. and Oroszlan, S. (1985) *Virology*, **145,** 280.
49. Gottlinger, H.G., Sodroski, J.G. and Haseltine, W.A. (1989) *Proc. Natl. Acad. Sci. USA*, **86,** 5781.
50. Peng, C., Chang, N.T. and Chang, T.W. (1991) *J. Virol.*, **65,** 2751.

51. Verma, I.M. (1975) *J. Virol.*, **15,** 843.
52. von der Helm, K. (1977) *Proc. Natl. Acad. Sci. USA*, **74,** 911.
53. Golomb, M. and Grandgenett, D.P. (1979) *J. Biol. Chem.*, **254,** 1606.
54. Golomb, M., Grandgenett, D.P. and Mason, W. (1981) *J. Virol.*, **38,** 548.
55. Grandgenett, D.P., Golomb, M. and Vora, A.C. (1980) *J. Virol.*, **33,** 264.
56. Grandgenett, D., Quinn, T., Hippenmeyer, P.J. and Oroszlan, S. (1985) *J. Biol. Chem.*, **260,** 8243.
57. Grandgenett, D.P., Knaus, R.J. and Hippenmeyer, P.J. (1983) *Virology*, **130,** 257.
58. Rho, H.M., Grandgenett, D.P. and Green, M. (1975) *J. Biol. Chem.*, **250,** 5278.
59. Grandgenett, D.P., Vora, A.C. and Schiff, R.D. (1978) *Virology*, **89,** 119.
60. Misra, T.K., Grandgenett, D.P. and Parson, J.T. (1982) *J. Virol.*, **44,** 330.
61. Grandgenett, D.P., Gerard, G.F. and Green, M. (1973) *Proc. Natl. Acad. Sci. USA*, **70,** 230.
62. Farmerie, W.G., Loeb, D.D., Casavant, N.C., Hutchison, C.A., Edgell, M.H. and Swanstrom, R. (1987) *Science*, **236,** 305.
63. Mous, J., Heimer, E.P. and Le Grice, S.F. (1988) *J. Virol.*, **62,** 1433.
64. Hansen, J., Schulze, T., Mellert, W. and Moelling, K. (1988) *EMBO J.*, **7,** 239.
65. Veronese, F.D., Copeland, T.D., DeVico, A.L., Rahman, R., Oroszlan, S., Gallo, R.C. and Sarngadharan, M.G. (1986) *Science*, **233,** 1289.
66. Restle, T., Müller, B. and Goody, R.S. (1990) *J. Biol. Chem.*, **265,** 8986.
67. Müller, B., Restle, T., Weiss, S., Gautel, M., Sczakiel, G. and Goody, R.S. (1989) *J. Biol. Chem.*, **264,** 13975.
68. Tisdale, M., Ertl, P., Larder, B.A., Purifoy, D.J., Darby, G. and Powell, K.L. (1988) *J. Virol.*, **62,** 3662.
69. Hizi, A., McGill, C. and Hughes, S.H. (1988) *Proc. Natl. Acad. Sci. USA*, **85,** 1218.
70. Prasad, V.R. and Goff, S.P. (1989) *Proc. Natl. Acad. Sci. USA*, **86,** 3104.
71. Johnson, M.S., McClure, M.A., Feng, D.F., Gray, J. and Doolittle, R.F. (1986) *Proc. Natl. Acad. Sci. USA*, **83,** 7648.
72. Hostomsky, Z., Hostomska, Z., Hudson, G.O., Moomaw, E.W. and Nodes, B.R. (1991) *Proc. Natl. Acad. Sci. USA*, **88,** 1148.
73. Davies, J.F., Hostomska, Z., Hostomsky, Z., Jordan, S.R. and Matthews, D. (1991) *Science*, **252,** 88.
74. DeVico, A.L., Copeland, T.D., Oroszlan, S., Gallo, R.C. and Sarngadharan, M.G. (1991) *J. Biol. Chem.*, 266, **6774.**
75. Lori, F., Scovassi, A.I., Zella, D., Achilli, G., Cattaneo, E., Casoli, C. and Bertazzoni, U. (1988) *AIDS Res. Hum. Retroviruses*, **4,** 393.
76. DeVico, A.L., Copeland, T.D., Veronese, F.D., Oroszlan, S., Gallo, R.C. and Sarngadharan, M.G. (1989) *AIDS Res. Hum. Retroviruses*, **5,** 51.
77. DeVico, A.L., Montelaro, R., Gallo, R.C. and Sarngadharan, M.G. (1991) *Virology*, **185,** 387.
78. North, T.W., Cronn, R.C., Remington, K.M., Tandberg, R.T. and Judd, R.C. (1990) *J. Biol. Chem.*, **265,** 5121.
79. Gilboa, E., Mitra, S.W., Goff, S. and Baltimore, D. (1979) *Cell*, **18,** 93.
80. Verma, I. M. (1977) *Biochim. Biophys. Acta*, **473,** 1.
81. Baltimore, D., Gilboa, E., Rothenberg, E. and Yoshimura, F. (1979) *Cold Spring Harb. Symp. Quant. Biol.*, **43** Pt 2, 869.
82. Rothenberg, E., Smotkin, D., Baltimore, D. and Weinberg, R.A. (1977) *Nature* (*Lond.*), **269,** 122.
83. Weinberg, R.A. (1977) *Biochim. Biophys. Acta*, **473,** 39.
84. Gilboa, E., Goff, S., Shields, A., Yoshimura, F., Mitra, S. and Baltimore, D. (1979) *Cell*, **16,** 863.
85. Rothenberg, E. and Baltimore, D. (1977) *J. Virol.*, **21,** 168.
86. Wang, L.H. (1978) *Annu. Rev. Microbiol.*, **32,** 561.
87. Bender, W. and Davidson, N. (1976) *Cell*, **7,** 595.
88. Kung, H.J., Hu, S., Bender, W., Bailey, J.M., Davidson, N., Nicolson, M.O. and McAllister, R.M. (1976) *Cell*, **7,** 609.
89. Panet, A., Baltimore, D. and Hanafusa, T. (1975) *J. Virol.*, **16,** 146.
90. Furuichi, Y., Shatkin, A.J., Stavenezer, E. and Bishop, J.M. (1975) *Nature* (*Lond.*), **257,** 618.
91. Taylor, J.M. (1977) *Biochim. Biophys. Acta*, **473,** 57.
92. Sawyer, R.C., Harada, F. and Dahlberg, J.E. (1974) *J. Virol.*, **13,** 1302.
93. Harada, F., Sawyer, R.C. and Dahlberg, J.E. (1975) *J. Biol Chem.*, **250,** 3478.
94. Peters, G., Harada, F., Dahlberg, J.E., Panet, A., Haseltine, W.A. and Baltimore, D. (1977) *J. Virol.*, **21,** 1031.
95. Harada, F., Peters, G.G. and Dahlberg, J.E. (1979) *J. Biol. Chem.*, **254,** 10979.
96. Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S. and Alizon, M. (1985) *Cell*, **40,** 9.

97. Cordell, B., Stavnezer, E., Friedrich, R., Bishop, J.M. and Goodman, H.M. (1976) *J. Virol.*, **19,** 548.
98. Peters, G. and Dahlberg, J.E. (1979) *J. Virol.*, **31,** 398.
99. Haseltine, W.A., Kleid, D.G., Panet, A., Rothenberg, E. and Baltimore, D. (1976) *J. Mol. Biol.*, **106,** 109.
100. Taylor, J.M. and Illmensee, R., (1975) *J. Virol.*, **16,** 553.
101. Verma, I.M., Meuth, N.L. and Baltimore, D. (1972) *J. Virol.*, **10,** 622.
102. Haseltine, W.A., Coffin, J.M. and Hageman, T.C. (1979) *J. Virol.*, **30,** 375.
103. Omer, C.A., Parsons, J.T. and Faras, A.J. (1981) *J. Virol.*, **38,** 398.
104. Swanstrom, R., Varmus, H.E. and Bishop, J.M. (1981) *J. Biol. Chem.*, **256,** 1115.
105. Haseltine, W.A., Maxam, A.M. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA*, **74,** 989.
106. Coffin, J.M., Hageman, T.C., Maxam, A.M. and Haseltine, W.A. (1978) *Cell*, **13,** 761.
107. Stoll, E., Billeter, M.A., Palmenberg, A. and Weissmann, C. (1977) *Cell*, **12,** 57.
108. Collett, M.S., Dierks, P., Parsons, J.T. and Faras, A.J. (1978) *Nature (Lond.)*, **272,** 181.
109. Luo, G.X. and Taylor, J. (1990) *J. Virol.*, **64,** 4321.
110. Friedrich, R. and Moelling, K. (1979) *J. Virol.*, **31,** 630.
111. Tanese, N. and Goff, S.P. (1988) *Proc. Natl. Acad. Sci. USA*, **85,** 1777.
112. Huber, H.E., McCoy, J.M., Seehra, J.S. and Richardson, C.C. (1989) *J. Biol. Chem.*, **264,** 4669.
113. Panganiban, A.T. and Fiore, D. (1988) *Science*, **241,** 1064.
114. Boone, L.R. and Skalka, A.M. (1981) *J. Virol.*, **37,** 109.
115. Rice, N.R. and Coggins, L. (1979) *J. Virol.*, **29,** 907.
116. Hu, W.S. and Temin, H.M. (1990) *Proc. Natl. Acad. Sci. USA*, **87,** 1556.
117. Goodrich, D.W. and Duesberg, P.H. (1990) *Proc. Natl. Acad. Sci. USA*, **87,** 2052.
118. Coffin, J.M. (1979) *J. Gen. Virol.*, **42,** 1.
119. Darlix, J.L., Bromley, P.A. and Spahr, P.F. (1977) *J. Virol.*, **23,** 659.
120. Olsen, J.C. and Watson, K.F. (1985) *J. Virol.*, **53,** 324.
121. Mitra, S.W., Godd, S., Gilboa, E. and Baltimore, D. (1979) *Proc. Natl. Acad. Sci. USA*, **76,** 4355.
122. Sorge, J. and Hughes, S.H. (1982) *J. Virol.*, **43,** 482.
123. Smith, J.K., Cywinski, A. and Taylor, J.M. (1984) *J. Virol.*, **49,** 200.
124. Finston, W.I. and Champoux, J.J. (1984) *J. Virol.*, **51,** 26.
125. Omer, C.A., Resnick, R. and Faras, A.J. (1984) *J. Virol.*, **50,** 465.
126. Champoux, J.J., Gilboa, E. and Baltimore, D. (1984) *J. Virol.*, **49,** 686.
127. Resnick, R., Omer, C.A. and Faras, A.J. (1984) *J. Virol.*, **51,** 813.
128. Luo, G.X., Sharmeen, L. and Taylor, J. (1990) *J. Virol.*, **64,** 592.
129. Huber, H.E. and Richardson, C.C. (1990) *J. Biol. Chem.*, **265,** 10565.
130. Wohrl, B.M. and Moelling, K. (1990) *Biochemistry*, **29,** 10141.
131. Rattray, A.J. and Champoux, J.J. (1987) *J. Virol.*, **61,** 2843.
132. Clavel, F. and Orenstein, J.M. (1990) *J. Virol.*, **64,** 5230.
133. Charneau, P. and Clavel, F. (1991) *J. Virol.*, **65,** 2415.
134. Roth, M.J., Schwartzberg, P.L. and Goff, S.P. (1989) *Cell*, **58,** 47.
135. Omer, C.A. and Faras, A.J. (1982) *Cell*, **30,** 797.
136. Boone, L.R. and Skalka, A.M. (1981) *J. Virol.*, **37,** 117.
137. Kung, H.J., Fung, Y.K., Majors, J.E., Bishop, J.M. and Varmus, H.E. (1981) *J. Virol.*, **37,** 127.
138. Stein, H. and Hausen, P. (1969) *Science*, **166,** 393.
139. Repaske, R., Hartley, J.W., Kavlick, M.F., O'Neill, R.R. and Austin, J.B. (1989) *J. Virol.*, **63,** 1460.
140. Papas, T.S. (1990) Editor. *Gene Regulation and AIDS*. Houston: Portfolio Publishing Co.
141. Molling, K., Bolognesi, D.P., Bauer, H., Büsen, W., Plassmann, H.W. and Hausen, P. (1971) *Nature New Biol.*, **234,** 240.
142. Baltimore, D. and Smoler, D. (1971) *Proc. Natl. Acad. Sci. USA*, **68,** 1507.
143. Keller, W. and Crouch, R. (1972) *Proc. Natl. Acad. Sci. USA*, **69,** 3360.
144. Grandgenett, D.P. and Green, M. (1974) *J. Biol. Chem.*, **249,** 5148.
145. Grandgenett, D.P., Gerard, G.F. and Green, M. (1972) *J. Virol.*, **10,** 1136.
146. Leis, J.P., Berkower, I. and Hurwitz, J. (1973) *Proc. Natl. Acad. Sci. USA*, **70,** 466.
147. Starnes, M.C. and Cheng, Y.C. (1989) *J. Biol. Chem.*, **264,** 7073.
148. Mizrahi, V. (1989) *Biochemistry*, **28,** 9088.
149. Oyama, F., Kikuchi, R., Crouch, R.J. and Uchida, T. (1989) *J. Biol. Chem.*, **264,** 18808.
150. Krug, M.S. and Berger, S.L. (1989) *Proc. Natl. Acad. Sci. USA*, **86,** 3539.
151. Schatz, O., Mous, J. and Le Grice, S.F. (1990) *EMBO J.*, **9,** 1171.
152. Destefano, J.J., Buiser, R.G., Mallaber, L., Myers, T.W., Bambana, R.A. and Fay, P.J. (1991) *J. Biol. Chem.*, **266,** 7423.

153. Taylor, J. and Sharmeen, L. (1987) *J. Cell Sci. Suppl.*, **7,** 189.
154. Champoux, J.J., Gilboa, E. and Baltimore, D. (1984) *J. Virol.*, **49,** 686.
155. Hizi, A., Hughes, S.H. and Shaharabany, M. (1990) *Virology*, **175,** 575.
156. Levin, J.G., Crouch, R.J., Post, K., Hu, S.C., McKelvin, D., Zweig, M., Court, D.L. and Gerwin, B.I. (1988) *J. Virol.*, **62,** 4376.
157. DeVico, A.L., Veronese, F.D., Lee, S.L., Gallo, R.C. and Sarngadharan, M.G. (1988) *AIDS Res. Hum. Retroviruses*, **4,** 17.
158. Wolfs, T.F., Geelen, J.L., Schellekens, H., Barin, F., Dekker, J.T. and Goudsmit, J. (1989) *AIDS Res. Hum. Retroviruses*, **5,** 535.
159. Laurence, J., Sanders, A. and Kulkosky, J. (1987) *Science*, **235,** 1501.
160. Sano, K., Lee, M.H., Morales, F., Nishanian, P., Fahey, J., Detels, R. and Imagawa, D.T. (1987) *J. Clin. Microbiol.*, **25,** 2415.
161. Ferns, R.B., Partriase, J.C., Tisdale, M., Hunt, N. and Tedder, R.S. (1991) *AIDS Res. Hum. Retroviruses*, **7,** 307.
162. Hollis, V.W., Jr., Aoki, T., Barrera, O., Oldstone, M.B. and Dixon, F.J. (1974) *J. Virol.*, **13,** 448.
163. Wuu, K.D., Graves, D.C. and Ferrer, J.F. (1977) *Cancer Res.*, **37,** 1438.
164. Jacquemin, P.C., Saxinger, C., Gallo, R.C., Hardy, W.D., Jr. and Essex, M. (1978) *Virology*, **91,** 472.
165. Walker, B.D., Blexner, C., Paradis, T.J., Fuller, T.C., Hirsch, M.S., Schooley, R.T. and Moss, B. (1988) *Science*, **240,** 64.
166. Steimer, K.S., Higgins, K.W., Powers, M.A., Stephans, J.C., Gyenes, A., George-Nascimento, C., Luciw, P.A., Barr, P.J., Hallewell, R.A. and Sanchez-Pescador, R. (1986) *J. Virol.*, **58,** 9.
167. Boucher, C.A., de Jager, M.H., Debouck, C., Epstein, L.G., de Wolf, F., Wolfs, T.F. and Goudsmit, J. (1989) *J. Clin. Microbiol.*, **27,** 1577.
168. Smoler, D., Molineux, I. and Baltimore, D. (1971) *J. Biol. Chem.*, **246,** 7697.
169. Baltimore, D. and Smoler, D. (1971) *Proc. Natl. Acad. Sci. USA*, **68,** 1507.
170. Goodman, N.C. and Spiegelman, S. (1971) *Proc. Natl. Acad. Sci. USA*, **68,** 2203.
171. Rho, H.M., Poiesz, B., Ruscetti, F. W. and Gallo, R.C. (1981) *Virology*, **112,** 355.
172. Watson, K.F., Schendel, P.L., Rosok, M.J. and Ramsey, L.R. (1979) *Biochemistry*, **18,** 3210.
173. Gregerson, D.S., Albert, J. and Reid, T.W. (1980) *Biochemistry*, **19,** 301.
174. Cheng, Y.C., Dutschman, G.E., Bastow, K.F., Sarngadharan, M.G. and Ting, R.C. (1987) *J. Biol. Chem.*, **262,** 2187.
175. Chandra, A., Gerber, T. and Chandra, P. (1986) *FEBS Lett.*, **197,** 84.
176. Garard, G.F., Rottman, F. and Green, M. (1974) *Biochemistry*, **13,** 1632.
177. Gerard, G.F. (1975) *Biochem. Biophys. Res. Commun.*, **63,** 706.
178. Robert-Guroff, M., Schrecker, A.W., Brinkman, B.J. and Gallo, R.C. (1977) *Biochemistry*, **16,** 2866.
179. Tamblyn, T.M. and Wells, R.D. (1975) *Biochemistry*, **14,** 1412.
180. Majumdar, C., Stein, C.A., Cohen, J.S., Broder, S. and Wilson, S.H. (1989) *Biochemistry*, **28,** 1340.
181. Majumdar, C., Abbotts, J., Broder, S. and Wilson, S.H. (1988) *J. Biol. Chem.*, **263,** 15657.
182. Kedar, P.S., Abbotts, J., Kovács, T., Lesiak, K., Torrence, P. and Wilson, S.H. (1990) *Biochemistry*, **29,** 3603.
183. Reardon, J.E., Furfine, E.S. and Cheng, N. (1991) *J. Biol. Chem.*, **266,** 14128.
184. Springgate, C.F., Battula, N. and Loeb, L.A. (1973) *Biochem. Biophys. Res. Commun.*, **52,** 401.
185. Sirover, M.A. and Loeb, L.A. (1974) *Biochem. Biophys. Res. Commun.*, **61,** 410.
186. Battula, N. and Loeb, L.A. (1975) *J. Biol. Chem.*, **250,** 4405.
187. Battula, N., Dube, D. and Loeb, L.A. (1975) *J. Biol. Chem.*, **250,** 8404.
188. Battula, N. and Loeb, L.A. (1976) *J. Biol. Chem.*, **251,** 982.
189. Roberts, J.D., Bebenek, K. and Kunkel, T.A. (1988) *Science*, **242,** 1171.
190. Preston, B.D., Poiesz, B.J. and Loeb, L.A. (1988) *Science*, **242,** 1168.
191. Takeuchi, Y., Nagumo, T. and Hoshino, H. (1988) *J. Virol.*, **62,** 3900.
192. Bebenek, K., Abbotts, J., Roberts, J.D., Wilson, S.H. and Kunkel, T.A. (1989) *J. Biol. Chem.*, **264,** 16948.
193. Roberts, J.D., Preston, B.D., Johnston, L.A., Soni, A., Loeb, L.A. and Kunkel, T.A. (1989) *Mol. Cell Biol.*, **9,** 469.
194. Kunkel, T.A. and Soni, A. (1988) *J. Biol. Chem.*, **263,** 14784.
195. Perrino, F.W., Preston, B.D., Sandell, L.L. and Loeb, L.A. (1989) *Proc. Natl. Acad. Sci. USA*, **86,** 8343.
196. Leider, J.M., Palese, P. and Smith, F.I. (1988) *J. Virol.*, **62,** 3084.
197. Dougherty, J.P. and Temin, H.M. (1988) *J. Virol.*, **62,** 2817.
198. Coffin, J.M. (1986) *Cell*, **46,** 1.

199. Parvin, J.D., Moscona, A., Pan, W.T., Leider, J.M. and Palese, P. (1986) *J. Virol.*, **59,** 377.
200. Hahn, B.H., Shaw, G.M., Taylor, M.E., Redfield, R.R., Markham, P.D., Salahuddin, S.Z., Wong-Staal, F., Gallo, R.C., Parks, E.S. and Parks, W.P. (1986) *Science*, **232,** 1548.
201. Yokoyama, S. and Gojobori, T. (1987) *J. Mol. Evol.*, **24,** 330.
202. Panet, A., Haseltine, W.A., Baltimore, D., Peters, G., Harada, F. and Dahlberg, J.E. (1975) *Proc. Natl. Acad. Sci. USA*, **72,** 2535.
203. Panet, A. and Berliner, H. (1978) *J. Virol.*, **26,** 214.
204. Haseltine, W.A., Panet, A., Smoler, D., Baltimore, D., Peters, G., Harada, F. and Dahlberg, J.E. (1977) *Biochemistry*, **16,** 3625.
205. Barat, C., Lullien, V., Schatz, O., Keith, G., Nugeyre, M.T., Grüninger-Leitch, F., Barre-Sinoussi, F., LeGrice, S.F. and Darlix, J.L. (1989) *EMBO J.*, **8,** 3279.
206. Colicelli, J. and Goff, S.P. (1986) *J. Virol.*, **57,** 37.
207. Baroudy, B.M., Fournier, M., Labouesse, J., Papas, T.S. and Chirikjian, J.G. (1977) *Proc. Natl. Acad. Sci. USA*, **74,** 1889.
208. Cordell, B., Swanstrom, R., Goodman, H.M. and Bishop, J.M. (1979) *J. Biol. Chem.*, **254,** 1866.
209. Baroudy, B.M. and Chirikjian, J.G. (1980) *Nucleic Acids Res.*, **8,** 57.
210. Araya, A., Keith, G., Fournier, M., Gandar, J.C., Labouesse, J. and Litvak, S. (1980) *Arch. Biochem. Biophys.*, **205,** 437.
211. Hu, J.C. and Dahlberg, J.E. (1983) *Nucleic Acids Res.*, **11,** 4823.
212. Garret, M., Romby, P., Giege, R. and Litvak, S. (1984) *Nucleic Acids Res.*, **12,** 2259.
213. Eiden, J.J., Quade, K. and Nichols, J.L. (1976) *Nature (Lond.)*, **259,** 245.
214. Araya, A., Hevia, E. and Litvak, S. (1980) *Nucleic Acids Res.*, **8,** 4009.
215. Araya, A., Sarih, L. and Litvak, S. (1979) *Nucleic Acids Res.*, **6,** 3831.
216. Sarih, L., Araya, A. and Litvak, S. (1988) *FEBS Lett.*, **230,** 61.
217. Levin, J.G. and Seidman, J.G. (1981) *J. Virol.*, **38,** 403.
218. Grandgenett, D.P., Vora, A.C. and Faras, A.J. (1976) *Virology*, **75,** 26.
219. Grandgenett, D.P. and Rho, H.M. (1975) *J. Virol.*, **15,** 526.
220. Faras, A.J., Garapin, A.C., Levinson, W.E., Bishop, J.M. and Goodman, H.M. (1973) *J. Virol.*, **12,** 334.
221. Sawyer, R.C. and Dahlberg, J.E. (1973) *J. Virol.*, **12,** 1226.
222. Peters, G. and Glover, C. (1980) *J. Virol.*, **35,** 31.
223. Panet, A. and Kra-Oz, Z. (1978) *Virology*, **89,** 95.
224. Peters, G.G. and Hu, J. (1980) *J. Virol.*, **36,** 692.
225. Waters, L.C. and Mullin, B.C. (1977) *Prog. Nucleic Acid Res. Mol. Biol.*, **20,** 131.