Reverse Transcriptase and Retroviral Replication

T. Matamoros, M. Álvarez, V. Barrioluengo, G. Betancor and L. Menéndez-Arias

Centro de Biología Molecular “Severo Ochoa” (Consejo Superior de Investigaciones Científicas – Universidad Autónoma de Madrid), Campus de Cantoblanco, Madrid Spain

1. Introduction

Within each viral particle, retroviruses package two copies of a single-stranded RNA genome of about 10 kb. All of the viral genomes contain three major genes, arranged in the order: 5’-gag–pol–env-3’, and some retroviruses may also have accessory genes (e.g. vif, vpr, tax, etc…). Structural proteins such as MA (matrix protein), CA (capsid protein) and NC (nucleocapsid protein) are encoded within gag. Envelope proteins that mediate viral entry (surface and transmembrane glycoproteins) derive from expression of the env gene. Virus-encoded enzymes such as the protease, the reverse transcriptase (RT) and the integrase, required to complete the viral life cycle, usually derive from the expression of pol. The reverse transcription of the viral single-stranded (+) RNA genome into double-stranded DNA is an essential step in retroviral replication and an important target for therapeutic intervention (for reviews, see Telesnitsky & Goff, 1997; Abbink & Berkhout, 2008; Sarafianos et al., 2009). Reverse transcription is a relatively complex process that requires the intervention of at least three elements: (i) the viral genomic RNA (that serves as template); (ii) a specific primer (i.e. a transfer RNA); and (iii) the viral RT. Retroviral RTs are enzymes that possess two activities: (i) a DNA polymerase activity that uses either RNA or DNA as template, and (ii) an RNase H activity, which degrades RNA from RNA/DNA hybrids. Unlike eukaryotic DNA polymerases, retroviral RTs are devoid of 3’→ 5’ exonucleolytic proofreading and show intrinsic error frequencies of around $10^{-4}$ to $10^{-5}$, well above the values reported for cellular DNA polymerases. Their lower accuracy together with their ability to switch templates during reverse transcription are major contributors to the extensive genetic variability observed in many retroviruses including human immunodeficiency virus type 1 and type 2 (HIV-1 and HIV-2). The diversification of retroviral genomes, based on the sequence of the pol gene (encoding for viral enzymes including RT) is illustrated in Fig. 1. Environmental factors as well as the molecular structure of retroviral RTs modulate their fidelity. In addition, retrovirus genetic variability can be affected by viral and cellular proteins. In this review, studies dealing with the molecular basis of fidelity of HIV-1 RT are summarized and discussed in the light of crystal structures of the enzyme. Structural information has been most useful in the design of antiretroviral drugs targeting the DNA polymerase activity of the RT. The last sections of this chapter summarize current
knowledge on the molecular basis of antiretroviral drug resistance and the mechanisms leading to selection of drug-resistant HIV.

Fig. 1. Dendrogram representing phylogenetic relationships between retroviruses. The unrooted neighbour joining dendrogram is based on pol sequences of seven retroviral genera (alpha-, beta-, gamma-, delta-, epsilon-, lenti- and spuma-like retroviruses). Reproduced from Jern et al. (2005); originally published by BioMed Central.

2. The process of reverse transcription

An overview of the reverse transcription process is shown in Fig. 2. Reverse transcription initiates after binding of a cellular tRNA primer (in HIV-1, trNA\textsuperscript{Lys,3}) to the primer binding site (PBS). The PBS is a sequence of 18 nucleotides, located downstream of the 5′-end of the genomic RNA. The viral RNA that serves as template for reverse transcription is flanked by repeat (R) sequences at its 5′ and 3′ termini. Upon annealing, the tRNA primer is extended.
Reverse Transcriptase and Retroviral Replication

up to the 5’ end of the genome, generating an intermediate which is known as the minus-strand strong-stop DNA ((-)ssDNA). Simultaneously, the RNase H activity of the RT degrades the RNA strand annealed to (-)ssDNA, and the (-)ssDNA is released to hybridize with the R sequence located at the 3’-end of the same RNA (intra-molecular jump) or with an R sequence located at a different viral genome (inter-molecular jump). This step is referred to as the first strand-transfer.

Fig. 2. Overview of the reverse transcription process.

After binding of the (-)ssDNA molecule to the R sequence, synthesis of the minus-strand DNA continues along the viral RNA together with simultaneous degradation of the template RNA. Polypurine tracts at the U3 region (3’-PPT) and in some retroviruses, at the centre of the genomic RNA (cPPT) resist degradation. These sequences serve as primers for synthesis of plus-strand DNA. Plus-strand DNA synthesis originating from the 3’-PPT continues to the 5’-end of the minus-strand until it reaches the 18th nucleotide in the tRNA where further synthesis is blocked by a methylated base. Priming from both PPTs involves discontinuous DNA synthesis that generates a 99-nucleotide DNA “flap” in the centre of the molecule. This product is referred to as plus strong-stop DNA ((+)ssDNA). The second strand transfer reaction involves the annealing of the (+)ssDNA to the 3’-end of the full-
length (-)strand DNA, through base pairing of the complementary PBS and PBS’ sequences. This strand transfer reaction is usually intra-molecular (Yu et al., 1998), and renders a circular intermediate.

The strand displacement activity of the RT is required to complete DNA synthesis and for the generation of a proviral DNA with duplicated long terminal repeats (LTR) at both ends. In the case of HIV-1 and other lentiviruses a “flap” appears in the centre of the DNA. This structure is eliminated by the activities of a cellular endonuclease that removes the “flap” and a ligase that joins the DNA ends to render the complete provirus that can eventually integrate in the host cell genome.

2.1 Role of the NC protein in reverse transcription

In most retroviruses, NC is a small nucleic-acid binding protein that derives from cleavage of the Gag polyprotein. In HIV-1, it has 55 amino acid residues and contains two conserved zinc fingers with the sequence $\text{CX}_{2}\text{CX}_{4}\text{HX}_{4}\text{C}$ (CCHC) connected by a short basic peptide linker. NC functions mainly as a nucleic acid chaperone (Levin et al., 2010), but it also contributes to packaging and dimerization of genomic RNAs, virus assembly, integration and reverse transcription.

NC influences almost every step in the reverse transcription process: (i) it facilitates annealing of the tRNA$_{\text{Lys,3}}$ to the viral genomic RNA; (ii) basic residues of NC bring the complementary tRNA and PBS sequences together, while its zinc fingers destabilize secondary structures in the 5’ LTR to prevent pausing of RT during reverse transcription; (iii) NC enhances minus strand transfer, by accelerating the annealing of 5’ repeat cDNA to the 3’ repeat sequence of the viral RNA genome; and (iv) NC increases RNase H activity, thereby enhancing cleavage of the donor template and promoting strand transfer by creating free cDNA available for interaction with the acceptor (Thomas & Gorelick, 2008). The ability of NC to destabilize secondary structures alleviates pausing and enhances RT processivity, leading to the generation of a greater proportion of full-length DNA products. NC is also important for removal of non-PPT RNAs, since it inhibits elongation of these primers without affecting extension of genuine polypurine tracts (Jacob & DeStefano, 2008).

2.2 Host and viral factors controlling reverse transcription

Interactions between RT and other viral proteins and/or cellular factors are likely to occur during early and late phases of viral replication. The formation of the reverse transcription complex involves packaging of tRNA in virions. This is possible because during assembly, Gag and Gag-Pol precursors interact with the aminoacyl-tRNA$_{\text{Lys,3}}$ synthetase (LysRS) which has bound tRNA$_{\text{Lys,3}}$ (reviewed in Abbink & Berkhout, 2008). RNA helicase A is a cellular protein that is able to rearrange RNA structures due to its unwinding activity on RNA secondary structures. RNA helicase A activity promotes viral reverse transcription by facilitating the accessibility of the RT to the viral RNA (Roy et al., 2006). Other cellular proteins such as HuR, AKAP149 and DNA topoisomerase I have been shown to interact with retroviral RTs (for a review, see Warren et al., 2009).

In addition, several factors have shown an influence on the integrity of the viral genome. Thus, in HIV-1, the accessory protein Vif (viral infectivity factor) promotes reverse transcription while increasing viral infectivity (Carr et al., 2008). In the absence of NC, Vif promotes annealing of tRNA$_{\text{Lys,3}}$, decreases pausing of the RT, destabilizes nucleic acid secondary structures, stimulates ssDNA synthesis and increases the efficiency of the first
strand transfer event during reverse transcription. Moreover, Vif is also an RNA chaperone. In the presence of NC, Vif inhibits NC-induced tRNA<sub>Lys,3</sub> annealing, RNA dimerization and reverse transcription initiation. Taken together, those results suggest that Vif could prevent premature initiation of reverse transcription (Henriet et al., 2007).

Through the expression of Vif, HIV-1 counteracts the antiviral effect of apolipoprotein B mRNA-editing, catalytic polypeptide enzymes (APOBEC3). APOBEC3 family members are cellular proteins with cytidine deaminase activity that have anti-HIV-1 activity (Sheehy et al., 2002; reviewed in Aguiar & Peterlin, 2008). APOBEC3F and APOBEC3G are encapsidated into budding virions. In the absence of Vif, APOBEC3F/G induce hypermutation of the HIV-1 genome. In vitro studies have also shown that in the presence of NC, APOBEC3G affects tRNA<sub>Lys,3</sub> annealing (Guo et al., 2007), and reduces strand transfer and integration (Mbisa et al., 2007).

2.3 RNase H activity and dynamics of RT/nucleic acid interactions

As mentioned earlier, retroviral RTs have an endonuclease activity that hydrolyzes the RNA strand in RNA/DNA hybrids to generate 5'-phosphate and 3'-hydroxyl ends. Retroviral RNase H has two major distinct modes of activity: (i) polymerase-dependent (DNA 3'-end-directed cleavage), and (ii) polymerase-independent (RNA 5'-end-directed cleavage). In addition, internal cleavage of RNA on RNA/DNA hybrids can also occur in the absence of 5’ or 3’ ends. The efficiency of this cleavage depends on the specific nucleotide sequence in the vicinity of the cleavage site (for a review, see Schultz & Champoux, 2008) (Fig. 3).

![Fig. 3. Retroviral RT binding modes to RNA/DNA hybrids and effects on RNase H cleavage.](www.intechopen.com)
In the polymerase-dependent mode, the polymerase active site of the RT locates at the 3'-end of DNA primer, while the RNase H domain is positioned 15-20 nucleotides away from the DNA primer terminus. In this situation, the RNase H makes primary cuts and then slides forward to make the secondary cuts 5-8 nucleotides from the 3'-end of DNA primer. These cleavages can occur during DNA synthesis and in the absence of dNTPs. Factors that decrease the rate of DNA polymerization result in an increased number of RNase H cuts, due to RT pausing. The distance from the recessed DNA 3’ and the nucleotide sequence in the vicinity of the cleavage site are important factors that contribute to the frequency of DNA 3´-end-directed cleavages (Basu et al., 2008; Schultz & Champoux, 2008).

Polymerase-independent cleavages depend on the precise position of the 5’-end of RNA fragments relative to the polymerase active site. In this mode, the polymerase domain of the RT binds the DNA strand in a site that is near the RNA 5’-end. The RNase H domain is positioned 13-19 nucleotides away from the 5’-end of the RNA. The distance from the recessed end, the nucleotide sequence in the vicinity of the cleavage site and the accessibility of the 5’-end (a gap of 2 or more bases is sufficient for such recognition) are important factors influencing RNA 5’-end-directed cleavages (Schultz & Champoux, 2008; Champoux & Schultz, 2009; Herschhorn & Hizi, 2010).

During reverse transcription, the polymerase-dependent mode is used during minus-strand DNA synthesis to cleave the RNA strand of the hybrid. However, the rate of DNA polymerization is about 10 times faster than the RNase H cleavage rate and therefore, insufficient for total RNA degradation and DNA release. Polymerase-independent RNase H activity is required to complete the process. Cleavages to generate the PPT primer are sequence-specific and can occur internally on RNA/DNA hybrids, without DNA synthesis. Polymerization-independent RNase H activity also appears to be important for removal of extended tRNA and PPT primers.

PPT primer removal could occur by an internal cleavage event at the RNA-DNA junction, or alternatively by an RNA 5’-end-directed cleavage after extension (Schultz & Champoux, 2008; Champoux & Schultz, 2009). However, molecular details underlying these events are largely unknown. Chemical probing, NMR spectroscopy and single-molecule fluorescence resonance energy transfer (FRET) have been recently used to explore molecular aspects involved in the communication between RT and the RNA/DNA hybrid (Liu et al., 2008; Abbondanzieri et al., 2008; reviewed in Fabris et al., 2009; Götte et al., 2010). These studies revealed that an important factor determining enzyme binding orientation is the backbone composition of nucleotides at the 5’-end of the primer. These nucleotides make specific contacts with RNase H primer grip residues, and thereby regulate the DNA polymerase activity of the RT (Abbondanzieri et al., 2008). These studies also showed that during reverse transcription, the RT could alternate binding modes (i.e. polymerase-dependent or independent) in the presence of PPT. The switching kinetics can be regulated. Thus, in the presence of dNTP, the polymerization binding mode predominates and flipping between both orientations decreases (Liu et al., 2008; reviewed in Götte et al., 2010).

2.4 Recombination and strand transfer
Recombination is the major source of genetic variability in retrovirus. Recombination can mediate the repair of defective retroviral genomes, increase viral diversity and accelerate the spread of beneficial mutations. In HIV-1, recombinant genetic forms derived from virus of
different subtypes emerged in individuals infected with viruses of two or more different subtypes. Inter-subtype recombination allowed the generation of around 50 different well-characterized circulating recombinant forms (CRFs) (for an update, see http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html).

Two relevant properties of retroviral reverse transcription contribute to the high frequency of recombination. First, retroviral genomic RNAs are encapsidated in pairs. Second, the replication machinery is prone to recombination, since template switching is required to complete proviral DNA synthesis (Yu et al, 1998; Ramírez et al., 2008; Onafuwa-Nuga & Telesnitsky, 2009). Recent studies have shown that the HIV-1 RT can stabilize short (2-nucleotide) duplexes of 3'-overhangs of the primer strand that are annealed to complementary dinucleotide tails of DNA or RNA template strands. This RT “clamping” activity that anneals RNA and DNA strands could have a role in strand transfer (Oz-Gleenberg et al., 2011).

3. The crystal structures of retroviral RTs

Although all retroviral RTs share similar activities and derive from expression of the viral pol gene, they differ in size and subunit composition (Herschhorn & Hizi, 2010). Lentiviral RTs are asymmetric heterodimeric enzymes, as shown for HIV-1 and equine infectious anemia virus (with subunits of 66 and 51 kDa), feline immunodeficiency virus RTs (67 and 54 kDa) and HIV-2 RT (68 and 55 kDa). Viral RTs arise from processing of the Gag-Pol precursor and contain DNA polymerase and RNase H domains in the large subunit and the DNA polymerase domain alone in the smaller subunit. The function of the small subunit is mainly structural. The bovine immunodeficiency virus RT has been expressed in bacteria as a 64/51-kDa heterodimer. A similar structure is shared by the human T cell lymphotropic virus type I RT that contains subunits of 62 and 49 kDa. The RTs of alpharetroviruses (e.g. avian sarcoma leukaemia virus) are larger heterodimers composed of subunits of 94 and 62 kDa, where the 94-kDa subunit also contains the viral integrase protein (Hizi & Joklik, 1977). On the other hand, the RT of the murine leukaemia virus (MLV) (a gammaretrovirus) is a monomer of 75 kDa, containing both DNA polymerase and RNase H domains. Less-studied recombinant RTs from mouse mammary tumour virus (66 kDa), bovine leukaemia virus (64-80 kDa) and prototype foamy virus (80 kDa) appear to be active as monomers or homodimers, although their subunit composition in the virion is unknown.

Our current knowledge on the structure of retroviral RTs is essentially based on many crystal structures of HIV-1 RT alone or in complex with nucleic acid or inhibitors, as well as additional information on a few crystal structures of HIV-2 RT and MLV RT. These studies have revealed the structure of the two catalytic domains of the RTs (DNA polymerase and RNase H), which are separated by a connection subdomain. As found in other DNA polymerases, RTs have a structure that resembles a right hand (Fig. 4), with three subdomains in the DNA polymerase domain, designated as fingers, palm and thumb.

3.1 HIV-1 RT

HIV-1 RT is a heterodimer composed of two subunits known as p66 (560 amino acids) and p51 (440 amino acids). The DNA polymerase active site residues (Asp110, Asp185 and Asp186) are located in the palm subdomain of p66 (Fig. 4). In p66, the palm and connection subdomains consist of five stranded β sheets with two α helices on one side, while the thumb subdomain is composed of a bundle of four helices (Kohlstaedt et al., 1992; Jacobo...
The fingers subdomain contains a mixed β sheet and three α helices. The RNase H domain consists of five β sheets flanked by four α helices (Fig. 4). The p66 and p51 subunits have similar folds but p51 is more tightly packaged. Fingers, palm, thumb and connection subdomains fold similarly in both subunits, but their spatial organization changes due to the different positioning of the fingers, thumb and palm subdomains. Both HIV-1 RT subunits form a large cleft, where the thumb subdomain of p51 and the connection subdomains of p66 and p51 form the “floor”, and fingers, palm and thumb subdomains of p66 provide lateral and apical interactions with the nucleic acid substrate (Jacobo-Molina et al., 1993). Active site residues in p66 are exposed to the cleft, but they are buried in the 51-kDa subunit (Rodgers et al., 1995).

Fig. 4. Crystal structure of HIV-1 RT. (A) Ribbon representation of the structure of HIV-1 RT showing the fingers, palm and thumb subdomains of p66 in blue, red and green, respectively, the connection subdomain in yellow, the RNase H domain in orange, and p51 in grey. (B) Ribbon representation of HIV-1 RT bound to a double stranded DNA. The template is shown in cyan and the primer in pale green. Catalytic Asp residues in the DNA polymerase domain are shown as yellow spheres and RNase H catalytic residues (Asp443, Glu478, Asp498 and Asp549) are represented as red spheres. (C) Individual structures of p66 (left) and p51 (right). Color codes are the same as in panels (A) and (B). Shown structures were all obtained with the PyMol molecular viewer (http://www.pymol.org) and Protein Data Bank coordinates 2HMI (Ding et al., 1998).

The structure of the binary complex of HIV-1 RT and double-stranded DNA (Jacobo-Molina et al., 1993; Ding et al., 1998) showed that the nucleic acid binding cleft can accommodate 17 nucleotides between the active sites of the DNA polymerase and the RNase H. Major
interactions between the RT and the template-primer are indicated in Table 1. The comparison of crystal structures of binary complexes with those obtained with unliganded RTs showed conformational changes involving the movement of the p66 thumb subdomain away from the fingers subdomain (Jacobo-Molina et al., 1993, Rodgers et al., 1995). In addition, the bound DNA adopts an A-type conformation in the vicinity of the DNA polymerase active site, but a B-like conformation near the RNase H domain. These changes in orientation involve a 40º bend of the DNA/DNA complex, near α-helix H in the thumb subdomain of p66 (Jacobo-Molina et al., 1993).

<table>
<thead>
<tr>
<th>Motif</th>
<th>RT subdomain</th>
<th>Residues involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primer grip</td>
<td>p66 palm (β12-β13)</td>
<td>227-235</td>
</tr>
<tr>
<td>Helix clamp</td>
<td>p66 thumb (αH-αl)</td>
<td>255-268 (αH)/278-286 (αl)</td>
</tr>
<tr>
<td>RNase H primer grip</td>
<td>p66 connection, p51 connection, p66 RNase H</td>
<td>358-361 (p66 connection), 395-396 (p51 connection), 473-476, 501, 505 (p66 RNase H)</td>
</tr>
<tr>
<td>Template grip</td>
<td>p66 fingers, p66 palm (β4, αB, β5, β8-αE)</td>
<td>73-77 (β4), 78-83 (αE), 86-90 (β5), 141-174 (β8-αE)</td>
</tr>
</tbody>
</table>

Table 1. HIV-1 RT residues and subdomains interacting with nucleic acid in the binary complex (Ding et al., 1998).

The crystal structure of a binary complex of HIV-1 RT and an RNA/DNA hybrid revealed only small differences in comparison with the HIV-1 RT/double-stranded DNA complex (Sarafianos et al., 2001). Thus, the distance between the DNA polymerase and RNase H active sites is slightly larger (18 nucleotides) in the RNA/DNA complex, with most of the contacts between RT and template-primer being maintained. However, the RNA/DNA hybrid makes more contacts with the p66 subunit of the RT at α-helix I (thumb subdomain), β-sheet 5 (palm subdomain) and with residues of the RNase H domain (Sarafianos et al., 2001; Tuske et al., 2004). Also, a number of contacts between the p51 subunit and the RNA template were not detected in the structure having the DNA/DNA substrate (Sarafianos et al., 2001). Nucleic acids have similar A-like/B-like conformations in complexes containing DNA/DNA or RNA/DNA. However, the transition from the A- to the B-forms generates a wider minor groove in the RNA/DNA complex, which together with additional contacts between the RNase H primer grip and the RNA template seem to be determinant for the RNase H catalytic activity.

An important milestone towards understanding the mechanisms and nucleotide specificity in DNA polymerization by retroviral RTs was the determination of the crystal structure of a ternary complex of HIV-1 RT bound to double-stranded DNA and an incoming dNTP (Huang et al., 1998). Nucleotide binding facilitates transition from an “open” conformation of the fingers subdomain in p66 (as observed in the structure of RT/DNA binary complexes) to a “closed” conformation where the β3-β4 hairpin loop in the fingers subdomain moves towards the p66 palm subdomain (see review by Sarafianos et al., 2009). This movement in the fingers subdomain is known to be the rate-limiting step in the polymerization reaction (Sarafianos et al., 2009), and brings amino acid residues Lys65 and Arg72 into close proximity with the incoming nucleotide (Tuske et al., 2004). Apart from these two residues, other important interactions in the nucleotide binding site are those established between the incoming dNTP and RT residues Asp113-Ala114-Tyr115-Phe116 and Gln151, as well as with the two divalent cations (probably Mg2+); and between RT residues Tyr183 and Met184 and
the DNA primer terminus (Huang et al., 1998). Binding of the incoming dNTP also produces a movement of the YMDD motif (including catalytic residues Asp185 and Asp186) that allows proper coordination of the catalytic aspartates with the metal cofactors, and triggers the nucleophilic attack of the 3’OH of the primer terminus on the α phosphorous of the incoming dNTP (Huang et al., 1998; Mendieta et al., 2008). This polymerization event renders an elongated DNA primer and a pyrophosphate molecule that is released in the reaction. Structural data suggest that the YMDD motif acts as a “springboard” supplying some of the energy required for translocation (Sarafianos et al., 2002).

3.2 HIV-2 RT
HIV-2 RT shares around 60% sequence identity with HIV-1 RT. Despite conservation of the cleavage site at the N-terminus of the RNase H domain, the size of the small HIV-2 RT subunit is uncertain (for a review, see Herschhorn & Hizi, 2010). There is only one crystal structure of HIV-2 RT available, and this was obtained as an unliganded form using the complete p68 subunit (559 residues) bound to a smaller subunit that contained only 427 residues due to degradation of the p68 polypeptide by bacterial proteases (Ren et al., 2002). The crystal structure of unliganded HIV-2 RT is similar to that of the HIV-1 RT, but the p68 thumb subdomain is rotated by 8º relative to the unliganded HIV-1 RT p66 subunit. HIV-2 RT heterodimers are also more stable than HIV-1 RT heterodimers (Divita et al., 1995).

3.3 MLV RT
Although MLV RT is quite different from HIV-1 and HIV-2 RTs (MLV RT is a monomer), the basic structure is the same: a right hand conformation with fingers, palm, thumb and connection subdomains and a C-terminal RNase H domain (Das & Georgiadis, 2004; Lim et al., 2006). Structural analysis revealed that HIV-1 and MLV RTs are rather different at their thumb and connection subdomains, while their fingers and palm subdomains show significant homology (for a review, see Coté & Roth, 2008). MLV RT is 111 amino acids longer than the p66 subunit of HIV-1 RT. Major differences are found at the N-terminus which is about 40 amino acids longer in MLV RT, and between the connection subdomain and the RNase H domain that contains 32 extra residues in the MLV RT. In the RNase H domain of this enzyme, there is a C-helix motif followed by a loop region of 11 amino acids, absent from HIV-1 RT. This structure is present in Escherichia coli RNase H.

4. The copying fidelity of RTs
The intrinsic fidelity of RTs has a major role in retroviral variability (for a recent review, see Menéndez-Arias, 2009), although cellular polymerases (e.g. replicative eukaryotic DNA polymerases α/β/δ/ε and RNA polymerase II) may also modulate the mutation rate in retroviruses. Other viral and host factors influencing the retroviral mutation rate include the cellular transcriptional machinery, physiological fluctuations of dNTP pools and asymmetric error repair. Several retroviruses (e.g. feline immunodeficiency virus, mouse mammary tumour virus, etc...) encode a dUTP pyrophosphatase (dUTPase) that prevents the incorporation of uracil into the viral genome. This protein is not present in HIV-1, but this virus contains a protein (Vpr) that allows for the encapsidation of isoforms of uracil DNA glycosylase. This cellular enzyme could contribute to reducing uracil content in the nascent viral DNA. In HIV, the virally encoded Vif protein overcomes the activity of APOBEC proteins, thereby contributing to maintain a non-lethal level of G→A mutations.
The intrinsic fidelity of purified retroviral RTs has been analyzed in vitro by using enzymatic (gel-based) or genetic assays (for a review, see Menéndez-Arias, 2002). Gel-based assays are based on the determination of kinetic parameters for the incorporation of correct and incorrect nucleotides on specific template-primers, and provide an estimate of the nucleotide selectivity of the DNA polymerase. The relevant kinetic parameters are the $k_{pol}$ (nucleotide incorporation rate constant) and the $K_d$ (equilibrium dissociation constant for dNTP) (Fig. 5). Their determination should be done under pre-steady-state conditions to avoid the contribution of the template-primer dissociation rate (Kellinger & Johnson, 2010). Since the fixation of a mutation involves nucleotide misincorporation followed by extension of the mismatched primer, similar assays should be carried out with template-primers having a mismatch at the 3' end of the DNA primer, in order to have a better estimate of the intrinsic fidelity of RTs. These assays have demonstrated that HIV-1 RT, MLV RT and RTs of other retroviruses are one to three orders of magnitude faster in extending mispaired template-primers than the dissociation of the retroviral RT from DNA. Reported mispair extension efficiencies for HIV-1 and MLV RTs are usually within the range of $10^{-2}$ to $10^{-4}$, while misinsertion ratios range from $10^{-3}$ to $10^{-6}$ (Menéndez-Arias, 2002; Matamoros et al., 2008; Álvarez et al., 2009).

$$k_{obs} = \frac{k_{pol} \times [dNTP]}{K_d + [dNTP]}$$

Fig. 5. Nucleotide incorporation reaction and example of the determination of kinetic parameters $k_{pol}$ and $K_d$ under pre-steady-state conditions. T/P$_n$ represents the template-primer with $n$ being the nucleotide length, and T/P$_{n+1}$ represents the same template-primer after the incorporation of one nucleotide. PPI stands for pyrophosphate.

Genetic assays are based on the expression of reporter genes. The most popular assay is probably the one based on the use as substrate of a gapped doubled-stranded M13mp2 DNA duplex that contains the $lacZ\alpha$ gene (Bebenek & Kunkel, 1995). A gap-filling reaction is carried out in the presence of RT and dNTPs, and mutants are identified after transformation of appropriate bacteria by the white/blue color of M13 plaques revealed using X-Gal indicator plates. Nucleotide sequence analysis of the M13 DNA obtained from white and pale blue plaques allows for the identification of base substitutions, insertions, deletions, etc… These assays have shown that the HIV-1 RT is 10 to 20 times less faithful than the MLV and avian myeloblastosis virus RTs (Roberts et al, 1989). Since the crystal structure of HIV-1 RT is known and this enzyme is an important target of antiretroviral therapy, there have been many studies on the effects of amino acid substitutions on the copying fidelity of this enzyme. A summary is presented in Table 2. Major structural determinants of the accuracy of DNA synthesis by the HIV-1 RT involve dNTP binding residues, amino acids that interact with the template-primer or the primer strand, minor groove binding track residues and amino acids located at the RNase H primer grip.
<table>
<thead>
<tr>
<th>RT location</th>
<th>Residue</th>
<th>Amino acid substitution</th>
<th>Effect on fidelity</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Enzymatic assays</td>
<td>Genetic</td>
</tr>
<tr>
<td>Nucleotide binding site</td>
<td>Lys65</td>
<td>K65R</td>
<td>▲</td>
<td>(8.1)</td>
</tr>
<tr>
<td>site residues</td>
<td>K65A</td>
<td>▲</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>Arg72</td>
<td>R72A</td>
<td>▼/▲</td>
<td>▲</td>
<td>(1.6)</td>
</tr>
<tr>
<td>Ala114</td>
<td>A114S</td>
<td>≈</td>
<td>≈</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A114G</td>
<td>≈</td>
<td>≈</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y115A</td>
<td>▼</td>
<td>▼</td>
<td>(4.0)</td>
</tr>
<tr>
<td></td>
<td>Y115F/V</td>
<td>▼</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y115L/I/N</td>
<td>▼</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y115M/H/G</td>
<td>▼</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y115C/S/W</td>
<td>▼</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q151M</td>
<td>≈</td>
<td>▲</td>
<td>(1.2)</td>
</tr>
<tr>
<td></td>
<td>Q151N</td>
<td>▲</td>
<td>▲</td>
<td>(13.1)</td>
</tr>
<tr>
<td></td>
<td>F160Y</td>
<td>≈</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F160W</td>
<td>≈</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M184I</td>
<td>▲</td>
<td>▲</td>
<td>(4.0)</td>
</tr>
<tr>
<td></td>
<td>M184A</td>
<td>▼</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M184V</td>
<td>▲/▼</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met184</td>
<td>M184L</td>
<td>▲</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>Located in the RT palm</td>
<td>Val148</td>
<td>V148I</td>
<td>ND</td>
<td>▲</td>
</tr>
<tr>
<td>subdomain, close to the</td>
<td>Ty183</td>
<td>Y183F</td>
<td>▲</td>
<td>▼</td>
</tr>
<tr>
<td>dNTP binding site</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resides that interact</td>
<td>Trp24</td>
<td>W24G</td>
<td>▲</td>
<td>ND</td>
</tr>
<tr>
<td>with the template strand</td>
<td>Phe61</td>
<td>F61A</td>
<td>ND</td>
<td>▲</td>
</tr>
<tr>
<td></td>
<td>F61G</td>
<td>≈</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L74V</td>
<td>▲</td>
<td>▲</td>
<td>(1.7-3.5)</td>
</tr>
<tr>
<td></td>
<td>V75A</td>
<td>▼</td>
<td>▼</td>
<td>(1.4)</td>
</tr>
<tr>
<td></td>
<td>V75F</td>
<td>▲</td>
<td>▲</td>
<td>(1.8)</td>
</tr>
<tr>
<td></td>
<td>V75I</td>
<td>▲</td>
<td>▲</td>
<td>(1.8-3.0)</td>
</tr>
<tr>
<td></td>
<td>D76V</td>
<td>▲</td>
<td>▲</td>
<td>(8.8)</td>
</tr>
<tr>
<td></td>
<td>R78A</td>
<td>ND</td>
<td>▲</td>
<td>(8.9)</td>
</tr>
<tr>
<td></td>
<td>E89G</td>
<td>▼/▲</td>
<td>▲</td>
<td>(1.4-2.0)</td>
</tr>
<tr>
<td></td>
<td>E89K</td>
<td>ND</td>
<td>▲</td>
<td>(1.2)</td>
</tr>
<tr>
<td></td>
<td>E89S</td>
<td>ND</td>
<td>▲</td>
<td>(1.6)</td>
</tr>
<tr>
<td></td>
<td>E89V</td>
<td>ND</td>
<td>▲</td>
<td>(1.3)</td>
</tr>
<tr>
<td>Resides that interact</td>
<td>Phe227</td>
<td>F227A</td>
<td>▲</td>
<td>ND</td>
</tr>
<tr>
<td>with the primer strand</td>
<td>Trp229</td>
<td>W229A</td>
<td>▲</td>
<td>ND</td>
</tr>
<tr>
<td>(β2-β3 hairpin)</td>
<td>M230L</td>
<td>ND</td>
<td>▼</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M230I</td>
<td>▼</td>
<td>▼</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gly231</td>
<td>▼</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tyr232</td>
<td>Y232A</td>
<td>▼</td>
<td>ND</td>
</tr>
<tr>
<td>RT location</td>
<td>Residue</td>
<td>Amino acid substitution</td>
<td>Effect on fidelity</td>
<td>References</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>---------</td>
<td>-------------------------</td>
<td>--------------------</td>
<td>------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Enzymatic</td>
<td>Genetic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>assays</td>
<td>assays</td>
</tr>
<tr>
<td>Minor groove binding track residues</td>
<td>Gln258</td>
<td>Q258A</td>
<td>▼ (2.0)</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Gly262</td>
<td>G262A</td>
<td>▼ (4.1)</td>
<td>36, 37</td>
</tr>
<tr>
<td></td>
<td>Trp266</td>
<td>W266A</td>
<td>▼ (3.0)</td>
<td>36, 37</td>
</tr>
<tr>
<td></td>
<td>Gln269</td>
<td>Q269A</td>
<td>▼ (2.6)</td>
<td>36</td>
</tr>
<tr>
<td>α-Helices H and I</td>
<td>Asp256</td>
<td>D256A</td>
<td>ND</td>
<td>▼ (1.2)</td>
</tr>
<tr>
<td>(additional thumb subdomain residues)</td>
<td>Lys259</td>
<td>K259A</td>
<td>ND</td>
<td>▼ (1.5)</td>
</tr>
<tr>
<td></td>
<td>Leu260</td>
<td>L260A</td>
<td>ND</td>
<td>▼ (1.2)</td>
</tr>
<tr>
<td></td>
<td>Lys263</td>
<td>K263A</td>
<td>ND</td>
<td>▼ (1.5)</td>
</tr>
<tr>
<td></td>
<td>Arg277</td>
<td>R277A</td>
<td>≈ ▲ (1.1)</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Gln278</td>
<td>Q278A</td>
<td>≈ ▲ (1.2)</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Leu279</td>
<td>L279A</td>
<td>≈ ▲ (1.1)</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Cys280</td>
<td>C280A</td>
<td>≈ ▼ (1.9)</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Lys281</td>
<td>K281A</td>
<td>≈ ▲ (1.1)</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Leu282</td>
<td>L282A</td>
<td>≈ ▲ (1.3)</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Arg284</td>
<td>R284A</td>
<td>≈ ▼ (1.1)</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Gly285</td>
<td>G285A</td>
<td>≈ ▲ (1.1)</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Lys287</td>
<td>K287A</td>
<td>≈ ▲ (1.3)</td>
<td>38</td>
</tr>
</tbody>
</table>

Symbols: ▼ and ▲ indicate a decrease or an increase in RT’s accuracy, respectively, ≈ indicates that the fidelity of the enzyme is not affected by the mutation. Numbers between parentheses represent the fold-change in mutant frequency as determined with the M13mp2 lacZα forward mutation assay. ND, not determined. References: (1) Garforth et al., 2010; (2) Mansky et al., 2003; (3) Shah et al., 2000; (4) Lewis et al., 1999; (5) Cases-González & Menéndez-Arias, 2005; (6) Jonckheere et al., 2000; (7) Martín-Hernández et al., 1996; (8) Martín-Hernández et al., 1997; (9) Boyer & Hughes, 2000; (10) Gutiérrez-Rivas & Menéndez-Arias, 2001; (11) Rezende et al., 1998a; (12) Weiss et al., 2002; (13) Weiss et al., 2000; (14) Jamburuthugoda et al., 2005; (15) Gutiérrez-Rivas et al., 1999; (16) Rezende et al., 1998b; (17) Hsu et al., 1997; (18) Oude Essink et al., 1997; (19) Pandey et al., 1996; (20) Hamburg et al., 1998; (21) Drosopoulos & Prasad, 1998; (22) Wainberg et al., 1996; (23) Bakhanashvili et al., 1996; (24) Weiss et al., 2004; (25) Agopian et al., 2007; (26) Fisher & Prasad, 2002; (27) Rubinek et al., 1997; (28) Matamoros et al., 2008; (29) Álvarez et al., 2009; (30) Kim et al., 1998; (31) Kim et al., 1999; (32) Hamburg et al., 2006; (33) Wisniewski et al., 1999; (34) Cases-González & Menéndez-Arias, 2004; (35) Wöhrle et al., 1997; (36) Beard et al., 1994; (37) Bebenek et al., 1995; (38) Beard et al., 1996.

Table 2. Effect of amino acid substitutions on the copying fidelity of HIV-1 RT.

5. HIV-1 RT as a drug target

Despite the contribution of other RTs to our general understanding of their biochemistry and role in retroviral replication, HIV-1 RT has been the most widely studied enzyme due to its relevance as a target for the development of antiretroviral drugs. Around half of the currently licensed drugs for treatment of HIV-1 infection are RT inhibitors. These drugs can be classified into nucleoside RT inhibitors (NRTIs) (Fig. 6) and non-nucleoside RT inhibitors (NNRTIs) (Fig. 7).
5.1 Inhibitors of the DNA polymerase activity of HIV-1 RT

5.1.1 Nucleoside RT inhibitors (NRTIs)

Currently, there are seven NRTIs approved for treatment of HIV-1-infected patients, as well as a nucleotide analogue inhibitor, tenofovir disoproxil fumarate, which is a precursor of an acyclic nucleoside phosphonate known as tenofovir (Fig. 6). Newly developed NRTIs are currently at different stages of drug development (for recent reviews, see Jochmans, 2008; Menéndez-Arias, 2008; Cahn & Wainberg, 2010).

Fig. 6. Chemical structures of approved NRTIs: β-D-(+)-3'-azido-3'-deoxythymidine (zidovudine, AZT), β-D-(+)-2',3'-didehydro-2',3'-dideoxythymidine (stavudine, d4T), β-D-(+)-2',3'-dideoxynosine (didanosine, ddl), β-D-(+)-2',3'-dideoxycytidine (zalcitabine, ddC), β-L-(−)-2',3'-dideoxy-3'-thiacytidine, (lamivudine, 3TC), (-)-(1S,4R)-4-[2-amino-6-(cyclopentylamino)-9H-purin-9-yl]-2-cyclopentene-1-methanol (abacavir), β-L-(−)-2',3'-dideoxy-5-fluoro-3'-thiacytidine (emtricitabine, FTC), and R-9-(2-phosphonomethoxy-propyl)adenine disoproxil fumarate (tenofovir disoproxil fumarate).

NRTIs are prodrugs that inside the cell will be phosphorylated by cellular kinases to their triphosphate forms. The resulting triphosphate derivatives are used as substrates by HIV-1 RT and incorporated at the 3'-end of the elongating DNA primer. Once incorporated, NRTIs act as chain terminators, blocking further DNA synthesis since these inhibitors lack the catalytically essential 3’OH group in the ribose ring. The therapeutic efficacy of NRTIs is determined by the efficiency of the activation process by cellular kinases (for a review, see Bazzoli et al., 2010). The long-term use of those drugs has been associated with a number of clinically relevant toxicities such as hyperlactatemia and lactic acidosis, lipoatrophy, etc. (Nolan & Mallal, 2004; Hammond et al., 2010). Some of those adverse effects appear to be related to the low specificity of the NRTIs, which do not markedly discriminate between RTs from different origins (HIV-1, HIV-2, SIV, MLV, etc) or cellular DNA polymerases. Most of these toxic effects are related to the inhibition of mitochondrial DNA polymerase γ (reviewed in Kohler & Lewis, 2007).

5.1.2 Non-nucleoside RT inhibitors (NNRTIs)

There are five NNRTIs that are currently approved for treating HIV-1 infections: nevirapine, delavirdine, efavirenz, etravirine and rilpivirine (Fig. 7), and others are currently in clinical trials (for a recent review, see Jochmans, 2008). Unlike NRTIs, these inhibitors do not need
intracellular metabolic activation to inhibit the viral replication. The interaction of these compounds with the RT induces conformational changes that impact the catalytic activities of the enzyme (Sluis-Cremer et al., 2004).

**Fig. 7. Chemical structures of approved NNRTIs.**

NNRTIs inhibit RT through binding to a hydrophobic pocket, located in the palm subdomain of the p66 subunit, at a distance of around 10 Å from the DNA polymerase catalytic site. The hydrophobic pocket is lined by aromatic residues Tyr181, Tyr188, Phe227, Trp229 and Tyr232, hydrophobic amino acids Pro59, Leu100, Val106, Val179 Leu234, and Pro236, and hydrophilic residues Lys101, Lys103, Ser105, Asp132 and Glu224, all of them in p66. In addition, two residues in p51 (Ile135 and Glu138) contribute to NNRTI binding. In the crystal structures of unliganded RT, there is no NNRTI binding pocket, but this pocket is formed when the inhibitor binds to the RT (Das et al., 2005). These drugs act as non-competitive inhibitors, binding to an allosteric site and interfering with dNTP incorporation, probably by altering the conformation of residues in the vicinity of the active site (Xia et al., 2007), rather than having a direct effect on phosphodiester bond formation (Spence et al., 1995).

### 5.2 Molecular mechanisms of HIV-1 RT resistance

Currently, most HIV-infected patients are treated with so-called “highly active antiretroviral therapy” (HAART), which consists of combinations of three or more inhibitors usually acting on two or more targets (e.g. RT DNA polymerase, NNRTI binding pocket, viral protease, etc.). One of the most successful combinations is made of three RT inhibitors: tenofovir, emtricitabine and efavirenz (see Figs. 6 and 7 for chemical structures). HAART can efficiently block viral replication, but modest levels of replication can lead to development of drug resistance. Although drug-resistant variants may have an overall rate of replication lower than the wild-type viruses in the absence of drugs; compensatory mutations may emerge to improve the viral replication capacity of these viruses (for reviews, see Menéndez-Arias et al., 2003; Martínez-Picado & Martínez, 2008).

#### 5.2.1 Resistance to NRTIs

Zidovudine (AZT) was the first drug approved by the U.S. Food and Drug Administration (FDA), in 1987. Resistance to AZT was soon detected in patients under monotherapy for 6
months or longer (Larder et al., 1989a,b). The first AZT resistance mutations identified were M41L, D67N, K70R, L210W, T215F/Y and K219Q (Larder and Kemp, 1989). After the incorporation of new drugs into anti-HIV therapies, such as zalcitabine (dDC), didanosine (ddI) or stavudine (d4T), many other resistance mutations were identified. Interestingly, AZT resistance mutations were also found in patients under d4T therapy. These mutations were called “thymidine analogue resistance mutations” (TAMs). TAMs reside primarily in the palm and fingers subdomain of the RT, relatively far away from dNTP binding site. However, several NRTIs resistance mutations affect residues located at the nucleotide binding site (e.g. K65R, Q151M, M184V, etc.) or at residues that interact with the 5’ single-stranded template overhang (e.g. L74V, V75I, V75T, etc) (Huang et al., 1998). HIV-1 becomes resistant to NRTIs by two different mechanisms: (i) by increasing discrimination against the triphosphate forms of the drug, or (ii) by increasing the RT’s ability to remove the NRTIs from blocked DNA primers, through phosphorolysis mediated by ATP or pyrophosphate (PPi). These two mechanisms are not mutually exclusive.

i. Resistance by NRTI discrimination during DNA polymerization

Mutations that confer resistance through this mechanism reduce the incorporation of the NRTI-triphosphate while the RT retains the ability to incorporate natural dNTPs. In these cases, resistance is usually associated with a decrease of the catalytic efficiency ($k_{pol}/K_d$) for NRTI-triphosphate incorporation, through the loss of affinity for the inhibitor (increase of its $K_d$) and/or by a reduction in the NRTI incorporation rate (reduced $k_{pol}$). A number of single amino acid substitutions found in vivo were associated with this mechanism (Table 3). A representative example is M184V/I, which confers high-level resistance to lamivudine (3TC) and emtricitabine (FTC). Met184 is within the highly conserved YXDD motif, and its substitution to Val or Ile causes a steric hindrance between the side-chain of these residues and the sulfur atom present in the ribose ring of 3TC and FTC (Sarafianos et al., 1999; Kellinger & Johnson, 2010).

Combination therapies can select for mutational patterns that confer multidrug resistance. Thus, NRTI-resistant viruses containing amino acid changes A62V, V75I, F77L, F116Y and Q151M were initially observed in virus isolated from patients receiving AZT and ddI (Shirasaka et al., 1995). These mutants exhibited high-level resistance to AZT, ddI and d4T, and low-level resistance to abacavir, 3TC, FTC and tenofovir in phenotypic assays. This multidrug resistance pattern is known as the “Q151M pathway”. Gln151 interacts with the ribose moiety and the nitrogen base of the incoming dNTP (Huang et al., 1998). Q151M and the accompanying mutations affect the hydrogen bonding network between the nucleotide deoxyribose and the RT, while increasing the relevance of the interaction between the enzyme and the 3’OH of the nucleotide (Deval et al., 2002). This allows the mutant RT to better discriminate between normal dNTPs (with 3’-OH) and triphosphate derivatives of NRTIs (without 3’-OH).

ii. Excision-mediated resistance.

The NRTI excision reaction is one of the most important pathways towards the acquisition of resistance to antiretroviral drugs. Early biochemical studies revealed that nucleotide discrimination was not impaired in AZT-resistant HIV-1 RTs (e.g. in the quadruple mutant D67N/K70R/T215F/K219Q) (Carroll et al., 1994; Krebs et al., 1997). However, in the presence of a PPi donor, AZT-resistant RTs were able to remove the inhibitor from the 3’-end of the primer, thereby allowing DNA synthesis to proceed through phosphorolysis.
<table>
<thead>
<tr>
<th>Mutation</th>
<th>Subdomain</th>
<th>Resistance</th>
<th>Mechanism</th>
<th>Comments</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>K65R</td>
<td>fingers</td>
<td>ddA, 3TC, FTC, tenofovir, ddI, abacavir</td>
<td>$k_{pol}$</td>
<td>Suppresses AZT resistance.</td>
<td>1-6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ddC</td>
<td>$k_{pol}K_d$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K70E</td>
<td>fingers</td>
<td>Tenofovir</td>
<td>$k_{pol}$</td>
<td>Mixed mechanism of resistance: discrimination and excision.</td>
<td>6, 7</td>
</tr>
<tr>
<td>L74V</td>
<td>fingers</td>
<td>Abacavir, ddI</td>
<td>$k_{pol}K_d$</td>
<td>Impairs viral replication. Antagonizes the effect of TAMs on AZT resistance.</td>
<td>2, 8, 9</td>
</tr>
<tr>
<td>V75I</td>
<td>fingers</td>
<td>acyclovir</td>
<td>$K_d$</td>
<td>Accessory mutation of the Q151 complex. Antagonizes the effect of TAMs on AZT resistance.</td>
<td>10-12</td>
</tr>
<tr>
<td>V75T</td>
<td>fingers</td>
<td>d4T</td>
<td>$K_d$</td>
<td>Mixed mechanism of resistance: discrimination and excision.</td>
<td>1, 13, 14</td>
</tr>
<tr>
<td>Q151M</td>
<td>palm</td>
<td>Multidrug-resistance</td>
<td>$k_{pol}$</td>
<td>Associated with 4 mutations: A62V, V75I, F77L and F116Y.</td>
<td>3, 15-18</td>
</tr>
<tr>
<td>M184V</td>
<td>palm</td>
<td>3TC, FTC, Abacavir</td>
<td>$k_{pol}K_d$</td>
<td>It can suppress AZT resistance.</td>
<td>2, 3, 19, 20</td>
</tr>
</tbody>
</table>

References: (1) Selmi et al., 2001; (2) Deval et al., 2004; (3) Deval et al., 2005; (4) Feng et al., 2006; (5) White et al., 2006; (6) Sluis-Cremer et al., 2007; (7) Kagan et al., 2007; (8) Miranda et al., 2005; (9) Frankel et al., 2005; (10) McMahon et al., 2008; (11) Matamoros et al., 2009; (12) Tchesnokov et al., 2009; (13) Petropoulos et al., 2000; (14) Lennerstrand et al., 2001; (15) Kaushik et al., 2000; (16) Deval et al., 2002; (17) Ray et al., 2002; (18) Frangeul et al., 2008; (19) Sarafianos et al., 1999; (20) Kellinger & Johnson, 2010. For additional references, see Menéndez-Arias, 2010. Acyclovir is a guanosine nucleoside analogue that inhibits herpes simplex virus replication. In patients co-infected with herpes simplex virus, it can inhibit HIV replication. Phosphorylated acyclovir inhibits HIV-1 RT, by terminating DNA chain elongation (Lisco et al., 2008).

Table 3. NRTI resistance mutations which increase the discrimination capacity. Mediated by ATP or PPI (Arion et al., 1998; Meyer et al., 1999; reviewed in Menéndez-Arias, 2008, 2010). The excision reaction involves a nucleophilic attack of the hydroxyl group.
substituent at the γ-phosphorous of ATP (the PPi donor) on the terminal phosphodiester bond of the blocked DNA primer (Matamoros et al., 2005). The efficiency of the excision reaction is influenced by several factors such as the nature of the template strand, the sequence context in which chain termination occurs, the Mg²⁺ concentration, or the specific nucleoside analogue that blocks the 3’ end of the primer. Thymidine analogues (AZT, d4T and ddT) and tenofovir are the best substrates of the reaction, while cytidine analogues (3TC and ddC) are removed very inefficiently (Mas et al., 2002; White et al., 2004; Marchand et al., 2007; Sluis-Cremer et al., 2007). Another important parameter for the excision reaction is the positioning of the 3’ end of the blocked primer. Newly incorporated NRTI, present at the primer site (P) (post-translocational complex) makes the terminal primer unavailable for excision. Nucleotide excision occurs if the 3’ end of the primer is located in the nucleotide binding site (N) (pre-translocational complex) (Fig. 8). The efficiency of this reaction also depends on the presence of a dNTP complementary to the next position on the template, that leads to the formation of a “dead-end complex” that blocks the excision reaction by forcing the equilibrium towards the post-translocational status (for a review, see Götte, 2006).

![Diagram](https://example.com/diagram.png)

Fig. 8. Model showing the equilibrium between AZTTP incorporation and AZT excision from chain-terminated primers. The inhibition of DNA polymerization is favoured when the AZT moiety is at the P site (in green) and the next dNTP complementary is able to occupy the N site (in purple) to form a “dead-end” complex. Resistance occurs when the PPi donor (usually ATP) is able to excise the inhibitor when this is positioned at the N site.

The recently solved crystal structures of complexes of wild-type and mutant M41L/D67N/K70R/T215Y/K219Q HIV-1 RTs bound to double-stranded DNA and the product of the excision reaction, AZT adenosine dinucleoside tetraphosphate (Ap₄AZT, Fig. 8) have revealed different orientations for the adenosine-monophosphate moiety of Ap₄AZT in the wild-type and mutant RT complexes (Tu et al., 2010). The structures are consistent
with an important role of π-π stacking interactions between the adenine ring of ApAZT and the side-chain of Tyr215, while phosphate portions of the excision product seem to be stabilized by the side-chain of Arg70 in the mutant enzyme. Secondary TAMs such as D67N and K219Q would enhance the ability of the adenosine-monophosphate moiety to interact with Tyr215, while contributing to the proper alignment of the phosphates during the excision reaction.

The emergence of HIV variants containing insertions or deletions in the fingers subdomain of the RT has been facilitated by the extensive use of NRTIs in the treatment of HIV-infected individuals (for review, see Menéndez-Arias et al., 2006). HIV-1 RT mutants with a dipeptide insertion (usually Ser–Ser, Ser–Gly or Ser–Ala) between codons 69 and 70 and additional TAMs such as M41L, A62V, T69S, K70R and T215Y showed very high levels of excision activity with primers terminated with AZT and d4T (Mas et al., 2000; Matamoros et al., 2004; Cases-González et al., 2007). In addition, deletions of residues at positions 67 or 69 of the RT have been associated with AZT resistance or hypersusceptibility (Imamichi et al., 2000; Villena et al., 2007; Kisic et al., 2008).

A number of mutations have been reported to reduce ATP-mediated excision activity in the presence of TAMs. Examples are K65R (White et al., 2006), L74V (Frankel et al., 2005; Miranda et al., 2005), V75I (Matamoros et al., 2009), Y181C (Selmi et al., 2003) or M184V (Götte et al., 2000; for a review, see Menéndez-Arias, 2008). Finally, mutations in the connection subdomain (for example, E312Q, G335C/D, N348I, A360I/V, V365I and A376S) or in the RNase H domain of RT (such a Q509L, H539N and D549N) can enhance resistance to AZT by altering the balance between excision and template RNA degradation (Nikolenko et al., 2005, 2007; Delviks-Frankenberry et al., 2007, 2008; Yap et al., 2007; Ehteshami et al., 2008), while thumb subdomain polymorphisms (P272A/R277K/T286A) could also modulate excision by affecting the interaction of the RT with the RNA/DNA hybrid, without involvement of the RNase H activity of the RT (Betancor et al., 2010).

5.2.2 Resistance to NNRTIs

More than 40 NNRTIs resistance mutations have been identified in vitro or in vivo (Tambuyzer et al., 2009; reviewed in Clotet et al., 2009). Nearly all of these mutations occur at the NNRTI binding site and lead to the loss of NNRTI binding affinity by the HIV-1 RT. In many cases, the rigid structure of the inhibitors decreases their ability to accommodate themselves inside the hydrophobic pocket of the mutated RT. This could be the result of a loss of key hydrophobic interactions. For example, mutations involved in resistance to delavirdine or nevirapine (e.g. Y181C, Y188L or F227L) affect interactions between aromatic residues in the RT and the NNRTIs. Rigid molecules such as nevirapine or efavirenz are prone to development of NNRTI resistance through single amino acid substitutions (e.g. K103N). Novel inhibitors (e.g. etravirine) have been designed as more flexible molecules that are more resilient to resistance since their binding to the RT is compromised only when two or more resistance mutations appear (for recent reviews, see Sluis-Cremer and Tachedjian, 2008; Sarafianos et al., 2009; de Béthune, 2010). Mutations outside the NNRTI binding site (e.g. N348I or A376S in the connection subdomain) have been recently associated with NNRTI resistance (Nikolenko et al., 2007; Yap et al., 2007; Schuckmann et al., 2010; Paredes et al., 2011). Their mechanism of action is indirect and relates to their effects on RT dimerization and interaction with the template-primer.
5.3 Inhibition of the RNase H activity of HIV-1 RT

RNase H activity offers an additional target for the development of antiretroviral drugs due to its essential role in HIV-1 replication. Unfortunately, progress in the development of RNase H inhibitors has been relatively slow, due to problems of citotoxicity or limited cell uptake, as well as an interest in exploiting other viral targets. As of today, there are no drugs of this kind approved for clinical use. RNase H inhibitors that interfere with metal cofactor binding include N-hydroxymides, hydroxylated tropolones and diketo acids (for reviews, see Jochmans, 2008; Adamson & Freed, 2010; Beilhartz & Götte, 2010). The interference with metal-ion coordination has been demonstrated by the structure of HIV-1 RT complexed with β-thujaplicinol (Himmel et al., 2009), a hydroxylated tropolone. Other RNase H inhibitors such as napthoquinones, vinologous ureas, small nucleic acid fragments (aptamers) and hydrazones show different mechanisms of action (Chung et al., 2010; Gong et al., 2011; reviewed in Jochmans, 2008; Beilhartz & Götte, 2010). For example, the crystal structure of HIV-1 RT complexed with an N-acyl hydrazone analogue shows that this drug binds more than 50 Å away from the RNase H domain, at a novel site between the polymerase active site and the NNRTI binding pocket (Himmel et al., 2006). This study suggests that these compounds may alter the trajectory of the template-primer inducing structural changes at the DNA polymerase primer grip and at the thumb subdomain, thereby preventing RNase H-catalyzed cleavage.

6. Conclusions

The discovery of reverse transcription forty years ago constituted a major breakthrough leading to the birth of modern retrovirology. With the identification of pathogenic human retroviruses and most notably HIV-1, the study of reverse transcription has devoted a lot of attention. These investigations have led to the development and clinical use of about a dozen compounds targeting HIV-1 RT that play a major role in the control of the disease. However, further studies should help us to find other targets in the reverse transcription process. Novel inhibitors targeting RNase H activity would be a valuable addition to the antiretroviral drug armamentarium, and experimental therapeutic approaches such as viral decay acceleration induced by mutagenic nucleosides could be promising if they prove to be efficient in vivo. In addition, there are important aspects related to retroviral replication that should be further studied. Namely, the contribution of viral and cellular proteins to the reverse transcription process; how reverse transcription is regulated inside the cell (or in virions); how cellular factors associate and control the reverse transcription complex after entering the cell and prior to proviral DNA synthesis; and also the dynamics of strand transfer that seem to be basic for understanding recombination in retroviruses.

7. Acknowledgments

This work was supported by grants of the Ministry of Science and Innovation of Spain (BIO2010/15542), Fundación para la Investigación y Prevención del SIDA en España (FIPSE) (grant 36771/08), Fondo de Investigación Sanitaria (through the “Red Temática de Investigación Cooperativa en SIDA” RD06/006), and an institutional grant from the Fundación Ramón Areces.
8. References


Rubinek, T., Bakhashvili, M., Taube, R., Avidan, O. & Hizi, A. (1997). The fidelity of 3’ misinsertion and mispair extension during RNA synthesis exhibited by two drug-


Since the discovery of the DNA structure, researchers have been highly interested in the molecular basis of genome inheritance. This book covers a wide range of aspects and issues related to the field of DNA replication. The association between genome replication, repair, and recombination is also addressed, as well as summaries of recent work of the replication cycles of prokaryotic and eukaryotic viruses. The reader will gain an overview of our current understanding of DNA replication and related cellular processes, and useful resources for further reading.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: