Summer 1994

Structure-Substrate Binding Relationships of HIV-1 Reverse Transcriptase

Steve Chien-Wen Huang
Old Dominion University

Structure-Substrate Binding Relationships of HIV-1 Reverse Transcriptase

by

Steve Chien-Wen Huang

B.S. May 1983, Chung-Hsing University

M.S. May 1985, Chung-Hsing University

A Dissertation Submitted to the Faculty of Old Dominion University and Eastern Virginia Medical School in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

OLD DOMINION UNIVERSITY
and
EASTERN VIRGINIA MEDICAL SCHOOL

August 1994

Approved by :

_____
Dr. Laura K. Moen, Director

_____
Dr. Frank J. Castora

_____
Dr. Christopher J. Osgood

_____
Dr. Mark S. Elliott

ABSTRACT

STRUCTURE-SUBSTRATE BINDING RELATIONSHIPS
OF HIV-1 REVERSE TRANSCRIPTASE

Steve Chien-Wen Huang

Old Dominion University
and
Eastern Virginia Medical School
1994

Director : Dr. Laura K. Moen

Human Immunodeficiency Virus, type 1 (HIV-1), is the causative agent of the Acquired Immunodeficiency Syndrome (AIDS). HIV-1 reverse transcriptase (RT), a heterodimer p66/p51, has been the major target for treatment of AIDS. The significance of the p51 subunit and the RNase H domain of p66 in terms of their influence on the RNA-dependent DNA synthesis was investigated. Clones of the wildtype HIV-1 RT subunits, p66 and p51, and a recombinant C-terminal deletion mutant, p64, [Barr, P. J. (1987) Bio/Technology **5**, 486-489] were employed to study the structure-substrate binding relationships of HIV-1 RT. The activity assays of RNA-dependent DNA synthesis on both poly(rA)(dT) and a random base RNA template hybridized with a DNA oligomer showed that p51 significantly affects the enzyme activity. The increase in processivity by p51 in the p66/p51 heterodimer was also demonstrated. These observations suggested that the integrity of p51 is important in subunit-interactions for maintaining a favorable conformation of the enzyme for optimal function. C-terminal deletion in p66 was seen to decrease the processivity. The dissociation constant (K_d) for poly(rA)(dT) obtained by nitrocellulose binding assays suggested that the processivity of HIV-

1 RT on poly(rA)(dT) correlated with the affinity for the substrate. The processivity of RT on RNA335-DNA20 was seen to be affected by the pause sites observed on the autoradiograms. The pauses of DNA synthesis tended to occur at positions of template containing poly G-C sequences. The order of processivity observed on RNA335-DNA20 was p64/p64, p66/p66 < p64/p51 < p66/p51. The C-terminal deletion in p66 was shown to affect the ability to extend the DNA strand on RNA template. In those non-wildtype forms of HIV-1 RT (p66/p66, p64/p64, and p64/p51), the affinity for primer-template seemed to be sensitive to the structure of the RNA template as seen when comparing K_ds between poly(rA)(dT) and RNA335-DNA20. The wildtype enzyme, p66/p51, appeared to have a similar affinity for both substrates.

Dedicated with love

to Mom :

Sue-Hwei Lee Huang

ACKNOWLEDGEMENTS

I wish to thank Dr. Laura Moen, my dissertation advisor, for her direction and encouragement throughout the entire course of this study.

I would also like to thank the other members of my dissertation committee, Dr. Mark Elliott, Dr. Frank Castora and Dr. Christopher Osgood, for their helpful suggestions, comments and advice on my dissertation.

Thanks are also given to Tom, Jackie and Rana for their friendship and help.

Also, I would like to express my appreciation to Chuck Huang Boyd for his friendship and proofreading the manuscript.

Finally, I am very grateful to my family for their constant love, encouragement and support during the past years.

TABLE OF CONTENTS

PAGE

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

I. Introduction

Human Immunodeficiency Virus, type 1 (HIV-1), is thought to be the causative agent of the Acquired Immunodeficiency Syndrome (AIDS) (1-5). HIV-1 reverse transcriptase (RT), an important viral enzyme responsible for converting the viral RNA genome into viral DNA after infection, has been the major target for treatment of AIDS (6-19). Although many inhibitors are known, such as 3'-azido-3'deoxythymidine (AZT), ddC and ddI which are dideoxymononucleoside analogs that cause chain termination during DNA polymerization, the use of these drugs is compounded by serious side effects such as bone marrow suppression, resumption of virus replication during therapy, and emergence of AZT-resistant HIV-1 mutants (20-22). Deeper knowledge regarding the molecular structure of this enzyme along with its enzymatic mechanism is highly desirable to develop a better drug design.

History of HIV (1-5)

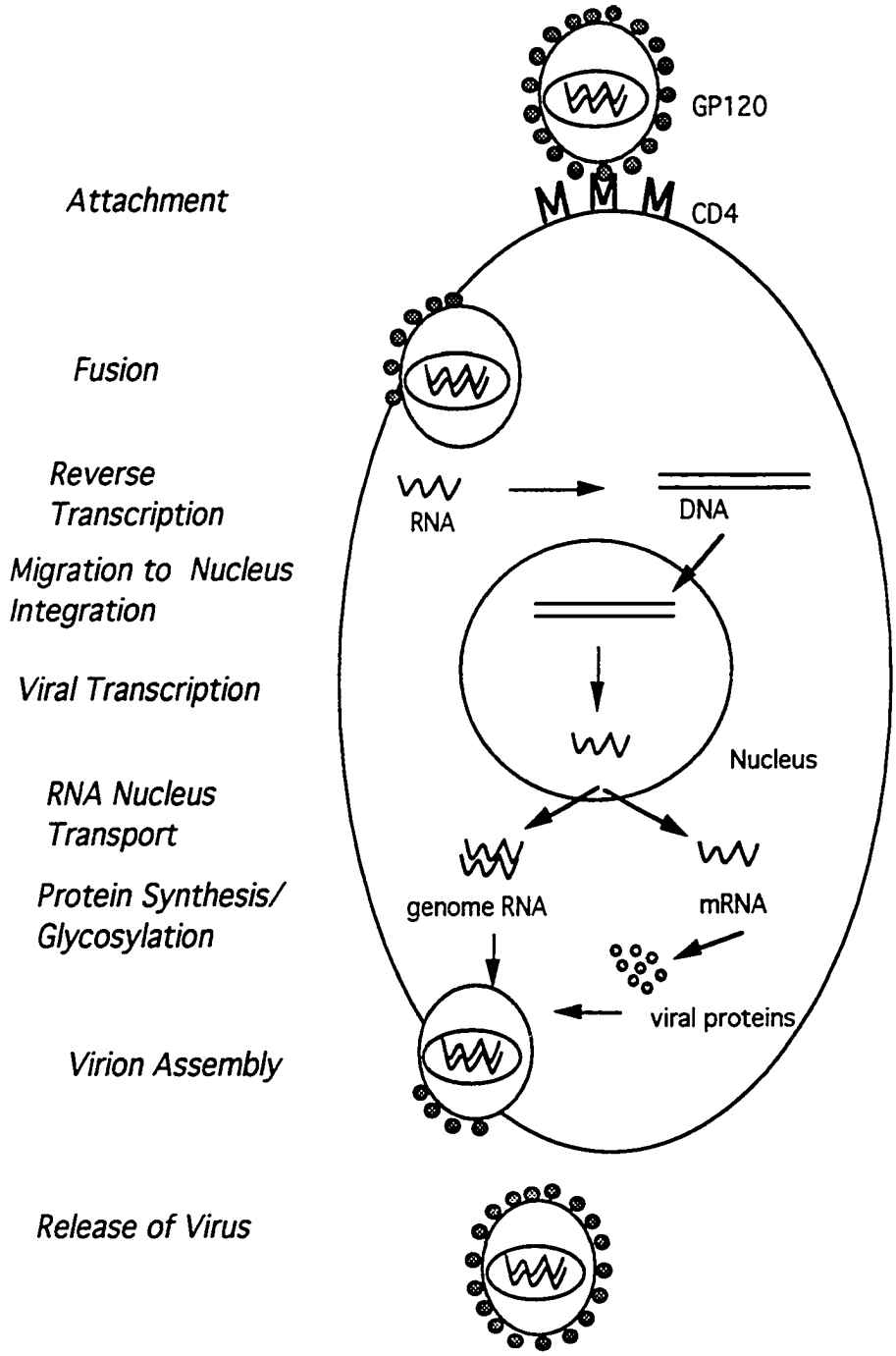
AIDS, an infectious disease spread by blood, by sexual intercourse, and from mother to child, has become an epidemic since the first AIDS cases were diagnosed in 1981 among young homosexual men in the U.S. HIV was discovered to be the causative agent during the two years from mid-1982 to mid-1984 (5). It was first observed that all AIDS victims suffered from a depletion of T4 cells, a specific subset of T cells. And it was suspected that a species of retrovirus, like HTLV-I which can infect T lymphocytes and cause highly

malignant cancer, was the pathogen. The evidence started with the finding that the reverse transcriptase activity, a main characteristic of retroviruses, was detected in the swollen lymph node of a young homosexual (5). This pathogen was identified as a new retrovirus. This new virus was found to grow only in T4 cells and not in the related cells called T8. Electron micrographs of the new virus were different from those of HTLV-I. Also, a viral protein p25 (or p24) found in the AIDS pathogen was not present in HTLV-I. A blood test was eventually formulated to show that this new virus was present in almost all people with AIDS, in a variable proportion of people who had received blood contaminated by the virus, and in no healthy heterosexuals (5). It was then firmly established that the new virus is the cause of AIDS and the virus was named as Human Immunodeficiency Virus, HIV.

Life cycle of HIV

HIV was later found to infect not only T4 cells but also macrophages (23). HIV enters its target cells by interacting with the molecule called CD4 (23). CD4 has a significant role in the immune function of T4 lymphocytes and is also found in the smaller numbers on some macrophages, allowing them to be infected. The life cycle of HIV is shown in Figure 1. The genes of a retrovirus are encoded as RNA. The cycle begins when an HIV particle binds to the outside of a cell and injects its core (23). The core includes two identical strands of RNA as well as structural proteins and enzymes. One enzyme, called reverse transcriptase, is responsible for converting the viral genetic information into DNA. The enzyme first makes a single-stranded DNA copy of the viral RNA, destroys the original RNA, and then makes a second DNA copy using the first one as the template (23). The viral genetic information, now in the form of double-stranded DNA, migrates to the cell nucleus. A second viral enzyme, called integrase, splices the

Figure 1 Life cycle of HIV. HIV infection begins when a virion, or virus particle, binds to the outside of a susceptible cell by the interaction between viral GP120 and host CD4 proteins. The virion then fuses with the cell, injecting the core proteins and two strands of RNA. The double-stranded DNA, synthesized by viral reverse transcriptase on the RNA genome, migrates to the nucleus and is integrated into the cell's own DNA. The provirus can then remain latent, giving no sign of its presence. Alternatively, it can commandeer the cellular mechanism to transcribe its genes into RNA, some of which are translated into proteins on ribosome. The proteins and the RNA genome are then assembled into new virions that bud from the cell (23).



HIV genome into the host cell DNA. The viral DNA will be duplicated with the cell's own genes every time the cell divides (23). The expression of the viral genes is regulated by the nucleotide sequence in the so-called long terminal repeats (LTR's) (23). LTR's are stretches of DNA at the ends of the viral genome which can direct enzymes belonging to the host cell to copy the proviral DNA into RNA. Some of the RNA will provide the genetic material for a new generation of virus. Certain other RNA strands serve as the mRNA's that guide cellular machinery in producing the structural proteins and enzymes of the new virus. Two different protein molecules in a ratio of about 20 to one will be assembled with the RNA genome (23). The more abundant molecule (gag) is the precursor of the protein shell that will enclose the RNA and enzyme in the complete virions. The other molecule (gag-pol) is larger; it contains the same structural components but includes additional segments that will become the viral enzymes. The two proteins migrate to the periphery of the cell as they are produced, a fatty acid at the end of each one attaches it to the inside of the cell membrane. As these precursors aggregate, they bind to one another and form a spherical structure that bulges outward under the cell membrane. Two strands of the viral RNA are drawn into this nascent virion as it takes shape. One of the enzymes contained in the gag-pol precursor, the viral protease, cuts itself free of the protein chain and cleaves the reverse transcriptase, the integrase and the additional molecules of protease. It then cleaves the gag polypeptides and what is left of the gag-pol polypeptides into four segments each. Three of the segments collapse to form a bullet-shape core surrounding the RNA and enzymes but the remaining segments stay attached to the inside of the cell membrane. The viral envelope proteins, which jut from the membrane like a set of spikes are made and transported to the cell surface independently of the core proteins. Each spike is a complex of two or three identical units that in turn

consist of two components: gp120 and gp41. These glycoprotein complexes are crucial to HIV's ability to infect new cells. The virions will sweep up the cell membrane with envelope proteins to enclose the core and release from the host.

Current therapeutic methods of RT inhibition

HIV-1 RT is responsible for converting the viral RNA into DNA and hence is the major target for antiviral treatment. Most of the effective inhibitors of HIV-1 RT are nucleoside analogs, the 2',3'-dideoxynucleosides, including AZT (20-22). These dideoxynucleosides are missing the 3'-hydroxyl which is required for further deoxynucleoside monophosphate (dNMP) incorporation. It is observed that HIV-1 RT has a higher affinity for dideoxynucleotides than for normal substrates (22). They inhibit HIV-1 RT after they are metabolized in vivo to the appropriate triphosphates by terminating the DNA chain growth. Unfortunately, all of these compounds have toxicities that limit their usefulness in the treatment for AIDS. For example, the most frequent toxicity of AZT is bone marrow suppression. In addition, the emergence of the drug resistant strains is observed. Comparative nucleotide sequence analysis of the reverse transcriptase-coding region from AZT sensitive and resistant isolates reveals four common amino acid substitutions: Asp67Asn, Lys70Arg, Thr215Phe, and Lys219Gln (22).

A variety of benzodiazepine and dipyridodiazepine derivatives have been identified as potent antiviral agents inhibiting HIV-1 RT as non-nucleoside analog inhibitors (50, 68, 69). Most appear to bind to a site on HIV-1 RT near Tyr181 which is distinct from the substrate binding site. Unlike the dideoxynucleoside analogs that can inhibit a number of human and animal retroviral DNA polymerases, these two groups of compounds are highly specific to HIV-1 RT. Several drugs selected from these two groups are under preclinical or clinical

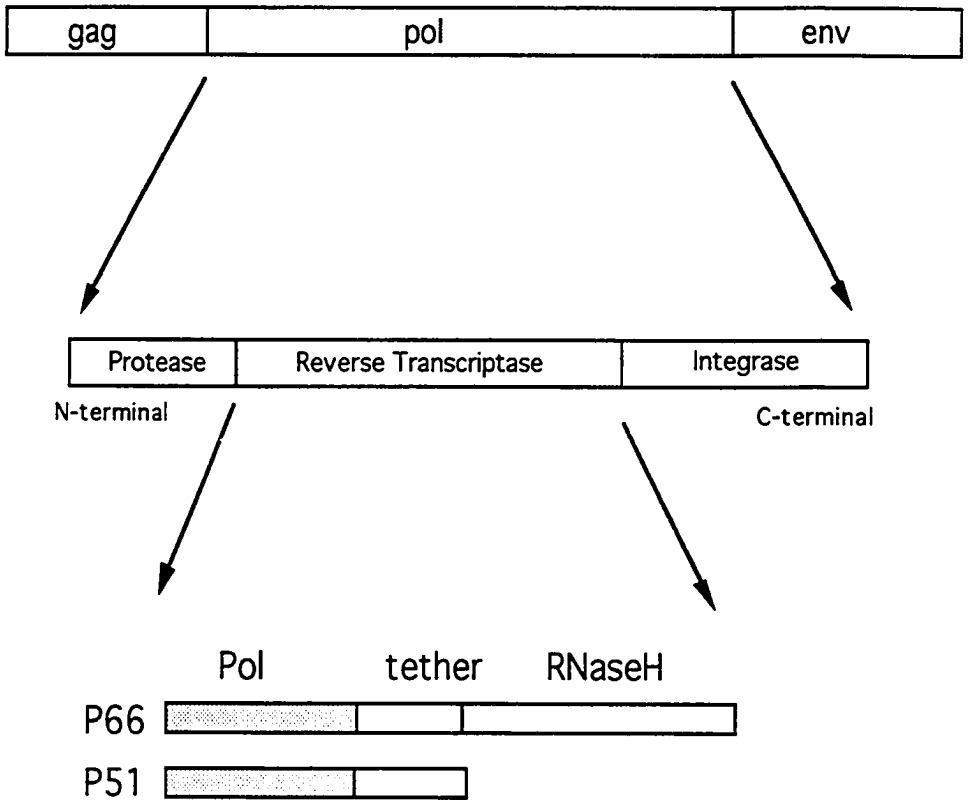
development. One of these non-nucleoside analogs, Nevirapine, is found to bind in a deep pocket that lies between beta sheet of the palm and the base of the thumb (50). It has been suggested that the binding of Nevirapine could prevent the movement of the thumb subdomain relative to the thumb subdomain during DNA synthesis or affect the conformation of the active site.

Despite intensive research efforts to identify compounds which inhibit HIV-1 RT, only AZT and ddI have been licensed to date for the treatment for HIV-1 infection. There is still great need to develop new drugs that have less cytotoxicity and are active against viral isolates resistant to AZT or ddI. Deeper knowledge about the structure and function relationships of HIV-1 RT is highly desirable for the development of more specific inhibitors.

Structure of HIV-1 RT

As shown in Figure 2, HIV-1 RT is encoded by *gag-pol* mRNA where the 5' end of the *pol* open reading frame overlaps the 3' end of *gag* with the *pol* frame by one nucleotide in the 5' direction (-1) (24). Ribosome frame-shifting is utilized to produce the gag-pol polypeptide which, in turn, gives rise to HIV-1 RT. In both virions and infected cells, HIV-1 RT has been characterized as a heterodimer consisting of two subunits: p66 and p51 (25-30). The two subunits have identical amino termini because p51 is derived by proteolytic processing of p66 by the viral protease with the cleavage site at Phe440-Tyr441 (31-35). It has been determined that the DNA polymerase and RNase H activities of HIV-1 RT map to the amino- and carboxyl- terminal portions connection region called the tether domain (28, 41, 46, 47). The p51 subunit lacks the RNase H domain. It has been reported that enzymatic activity is almost exclusively confined to the dimeric form of the enzyme and only one primer-template binding site is present in a dimer (48, 49). It has also been shown that the p66 subunit provides the

Figure 2 Schematic representation of the *pol* gene of HIV-1 RT and its relationship to protease, RT and integrase .

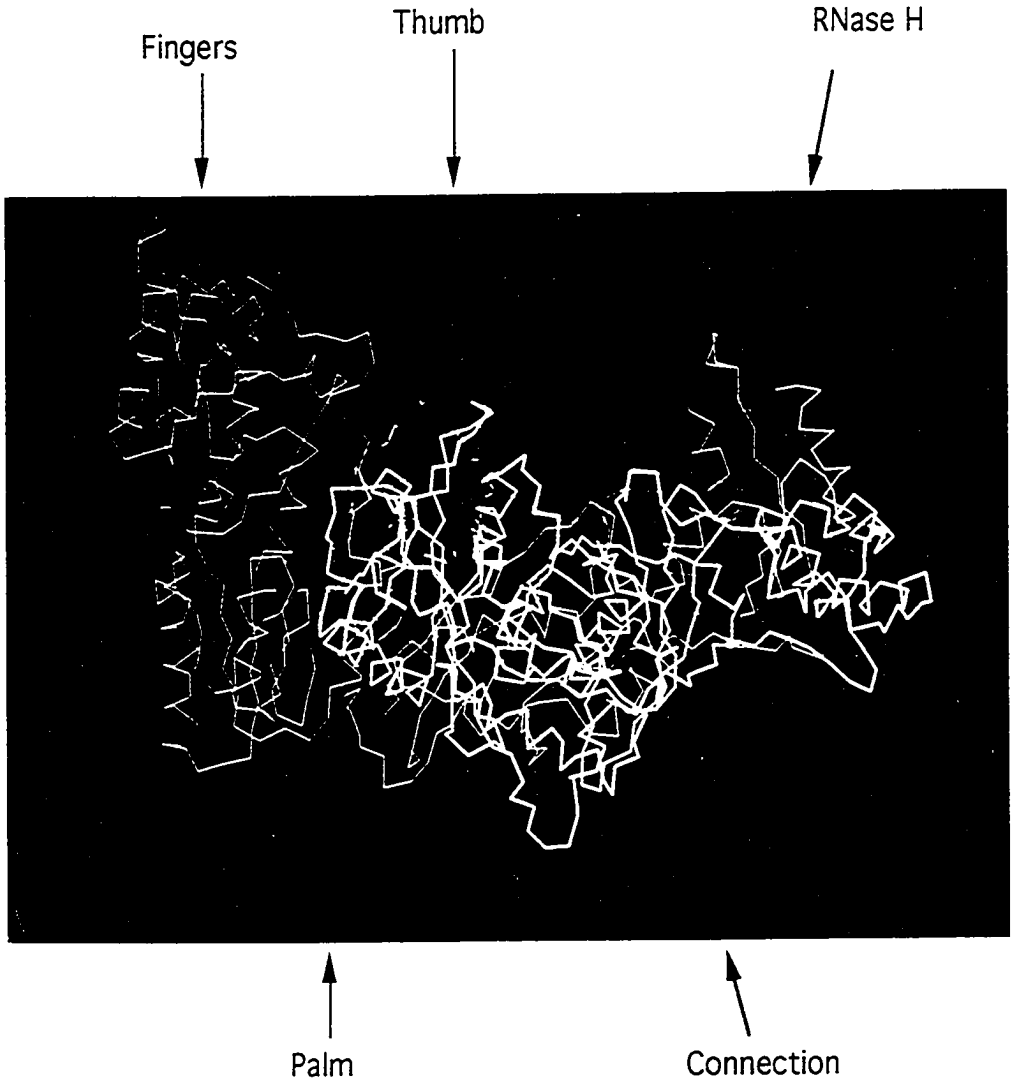


polymerization active site in the p66/p51 heterodimer (49). According to the crystal structure, the p66 subunit is folded into five separate regions: the RNase H domain, the connection domain, and three subdomains of the *pol* domain, namely fingers , palm, and thumb (50, Figure 3). The p66 subunit is highly elongated in folding, having dimensions of approximately 110 by 30 by 45 A^0, and a cleft is formed by the finger, palm, and thumb subdomains (50). The conformation of p51 is astonishingly different from that of p66, although both bear the same amino acid sequences (50). Unlike p66, p51 has no cleft, and residues thought to be involved in catalysis (Asp185, Asp186, Asp110) are essentially buried (50). The thumb domain is further from the fingers and palm subdomains. The connection domain lies within and fills the expanded cleft (50). The tip of the fingers and the connection domain of p51 interact with the palm and connection sudomains of p66 (50). The extended thumb of p51 contacts the RNase H domain of p66 (50). The crystal structure of HIV-1 RT implies that one of the p66 subunits of the p66/p66 homodimer should assume the conformation of p51 and unravel its RNase H domain for cleavage (50).

Function of HIV-1 RT

HIV-1 RT can polymerize a DNA strand from either an RNA or a DNA template. The enzyme also encodes an RNase H activity that hydrolyzes the RNA of RNA-DNA duplexes (51-57). As shown in Figure 4, the primer for minus strand DNA synthesis is a specific $tRNA^{(Lys)}$ molecule that is packaged in the virion and base paired to the primer-binding site (PBS) located near the 5' end of the RNA genome (58). Synthesis of the minus strand is initiated from the 3' end of the tRNA molecule and continued to the end of the viral genome. The hybridized RNA (R and U5 sequences) is removed and the short DNA minus strand is translocated to the 3' end of the same or a second RNA molecule,

Figure 3 A stereo α-carbon model of p66/p51. The p66 subunit is in blue color. The fingers, thumb, palm, connection, and RNase H subdomains are indicated by arrows. Residues 539 - 560 of p66 (the region deleted in p64) are shown in green. The p51 subunit is in white color. The double helix in red is the template-primer. Model was generated with SYBYL on a Silicon Graphics computer from the α-carbon coordinates reported by [Arnold, E. *et al.* (1992) Nature **357**, 85-89].



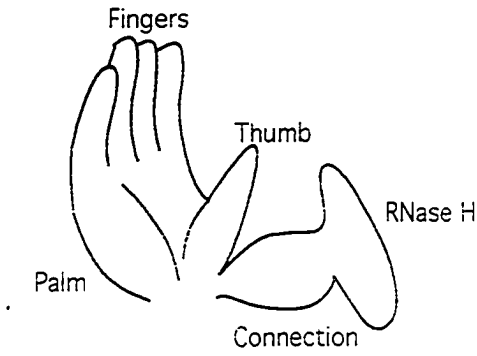
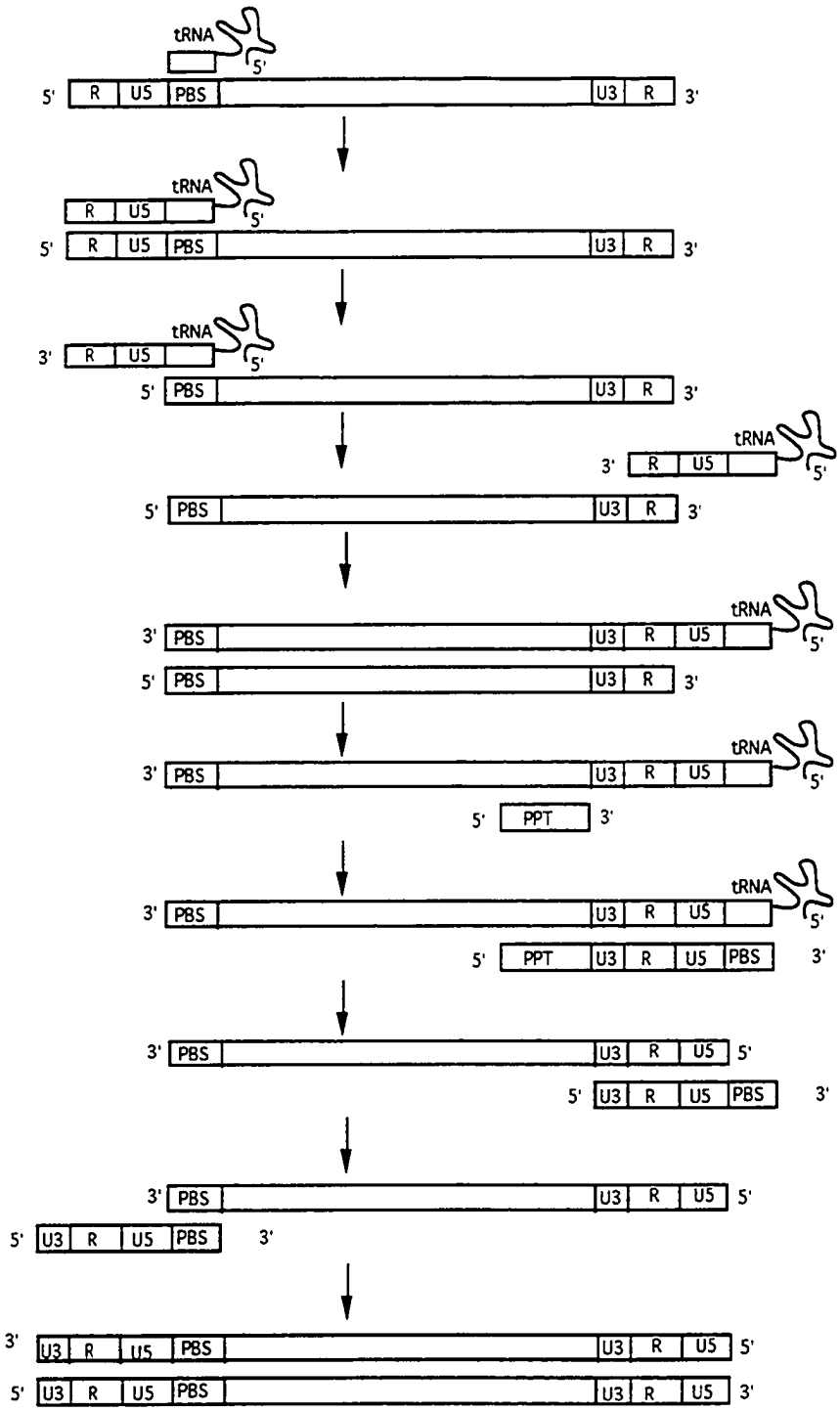


Figure 4 The process of reverse transcription by HIV-1 RT (58).

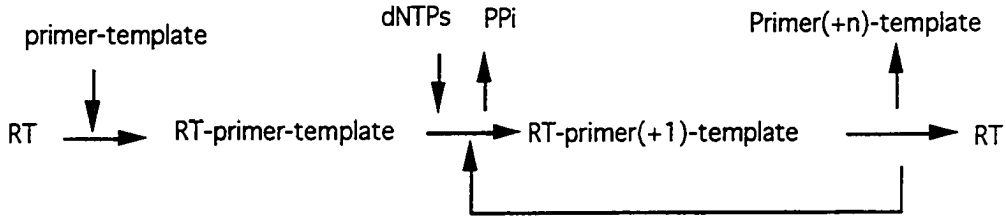


By reconstituting p66, p51, and p64, we have been able to use four forms of enzymes for structure-function study. One form, p66/p51, is the wildtype heterodimer. Another, p66/p66, is the homodimer which contains an extra RNase H domain (p15) in the non-catalytic subunit. The mutant heterodimer, p64/p51, has 22 amino acid residues deleted in the C-terminus and lacks the RNase H function, and p64/p64 is the mutant homodimer that contains C-terminal deletion at p66 and has an extra stretch of polypeptide in the non-catalytic subunit.

The activity assays of RNA-dependent DNA synthesis on both poly(rA)(dT) and RNA335-DNA20 demonstrated that the replacement of p51 for the non-catalytic subunit significantly increases the enzyme activity (Figure 12, 13, 20, and 21). The results showed that the conditions used for reconstitution in this study favor the formation of a more active enzyme, the heterodimer. This result was supported by the observation that the enzyme activity is exclusively confined to the dimer forms in an investigation of the process of dimer formation and the properties of the different mono- and dimeric forms of HIV-1 RT (28, 93). Also, it has been shown that individually purified p66 and p51 were capable of binding to form 1 : 1 heterodimer with an association constant 10-fold greater than that observed for p66/p66 homodimer formation (101).

The reconstituted heterodimer, in both wildtype and mutant, was shown to have better processivity on poly(rA)(dT) (Figures 14 -17). Since monomeric p66 has little activity (28, 93), these increases in enzyme activity and processivity suggested that the p51 polypeptide has an important role in the enzyme structure and function. One possible role could be that p51 improves the enzyme processivity through its contribution to a better primer-template interaction and/or a better conformation of the enzyme active site. The primer-template interaction with the enzyme was hence investigated.

Figure 5 DNA polymerase mechanism of HIV-1 RT. A bireactant biproduct mechanism has been proposed for DNA synthesis by HIV-1 RT. In this mechanism, binding of the reactants is ordered with the primer-template binding first followed by the binding of the dNTP that is complementary to the next appropriate template residue. The incorporation of dNMP into the nascent chain then occurs via an SN2 displacement of pyrophosphate (the first product) of the dNTP by the 3' hydroxyl at the primer terminus. Multiple dNMP additions can occur before the enzyme dissociates from the extended primer-template, the second product.



pyrophosphate (the first product) of the dNTP by the 3' hydroxyl at the primer terminus. Multiple dNMP additions can occur before the enzyme dissociates from the extended primer-template, the second product. HIV-1 RT is highly error-prone and produces approximately 5 to 10 nucleotide errors per genome every round of replication in vivo (67). This high error rate of HIV-1 RT is believed to be responsible for the hypermutability of HIV (67).

The RNase H domain of HIV-1 RT was assigned to the residues Glu348 through Ile559 of the polypeptide (47). Asp443, Glu478, Asp498, and Asp549 were suggested to be important in RNase H activity (98). The HIV-1 RT RNase H domain was found to display both endonuclease and 3' to 5' exonuclease activities (86). The distance between the RNase H and polymerase active sites was found to correspond to the length of a 15-16 nucleotide DNA-RNA heteroduplex (99). Also, it was found that the RNase H function is much less active than the polymerase function during processive DNA synthesis (87). The endonuclease function of RNase H domain was found to be in the direction of 3' to 5' and displayed a partially processive activity. The RNase H domain expressed in bacteria was found to be inactive by itself, and the addition of the p51 polypeptide led to enzymatic activity (84). This observation suggested the interdependence of polymerase and RNase H domains.

Interaction between HIV-1 RT and primer-template

In the crystal structure of HIV-1 RT, an obvious groove which exists between the RNase H domain and polymerase domain is presumably the primer-template binding site (50). DNA synthesis on an RNA template by HIV-1 RT is accompanied by the cleavage of the RNA template by RNase H; the RNase H cleavage site has been estimated to be some 15 to 16 nucleotides downstream

of the synthesis. Thus, the RNA-DNA duplex of the primer-template must lie between these two active sites with the 3' end of the RNA contacting the RNase H active site marked by the binding of the divalent cations in the crystal structure, and the 3' end of the DNA of the duplex in contact with the DNA polymerase active site. In a fitting model , the RNA-DNA can be fit snugly into the groove observed in the crystal structure when the DNA is modestly bent toward the protein. The cross-linking of dT_{15} to HIV-1 RT homodimers showed the cross-links at residues Leu289 to Thr290 and Leu295 to Thr296 which are in the thumb subdomain (69). It is also reported that the N-terminus of residues 1 to 230 of HIV-1 RT is capable of nucleic acid binding (70). The location of p51 relative to the RNA-DNA-HIV-1 RT fitting model suggests that the connection domain of p51 has a possible role in binding the tRNA-template complex by contacting the duplex as well as the eighteen anticodon and D stems and loops of tRNA (50).

Steady-state analyses have been performed on HIV-1 RT to determine the dissociation constant for primer-template. A summary of the measured dissociation constants of the wild-type HIV-1 RT (p66/p51) for primer-template is given in Table 1. Huber *et al.* (67) have demonstrated that HIV-1 RT forms a stable complex with poly(rA)(dT) primer-template in the absence of Mg^{++} and dTTP with a K_d of 3 nM. Krug *et al.* (63) have used globin mRNA hybridized to dT_{15} and demonstrated that K_d calculated by Scatchard method is 11 nM. The amount of globin mRNA-dT bound to the enzyme is obtained by extrapolating the burst of incorporation at zero time. Reardon's studies (71) have revealed a K_d of 10 nM for HIV-1 RT binding to a template-primer (RNA44-DNA21) by measuring the burst of product formation upon addition of dTTP. Kati *et al.* (72) have used rapid transient kinetic method with defined synthetic RNA45-DNA25 to obtain a K_d of 10 nM for the interaction between HIV-1 RT and this template-primer. Hsieh *et al.* (73), using a quadratic binding equation: $[E.D] = 0.5 ([E]_t +$

Table 1 Summary of the measured dissociation constants of HIV-1 RT (p66/p51) for primer-template.

Author	Huber	Krug	Reardon	Kati	Hsieh	Beard
primer-template	rAdT	mRNAdT	r44/d21	r45/d25	d32/d23	rAdT
Reference	67	63	71	72	73	74
K_d (nM)	3	11	10	26	15	< 1

$[D]_t + K_d) - 0.5 ([E]_t + [D]_t + K_d)^2 - 4 [E]_t [D]_t)^{1/2}$, have obtained a K_d of 26 nM for the interaction between HIV-1 RT and a DNA32-DNA23 template-primer. Binding of this 23/32-mer primer-template by HIV-1 RT was also examined by the gel mobility shift assay and results fit to a hyperbolic binding curve with K_d of 15 nM. Beard *et al.* (74) have used a template-primer challenge assay to quantitatively follow RT-template-primer complex formation and observed that HIV-1 RT (both p66/p66 and p66/p51) bind to poly $(rA)(dT)_{16}$ tightly with a $K_d < 1$ nM. As shown in Table 1 all the values are in nM range. Also, the K_d for poly $(rA)(dT)_n$ was found to be dependent on the length of the primer; short primer binding had a lower affinity.

Processivity and pausing of HIV-1 RT DNA synthesis

The processivity of a polymerase is defined as the number of nucleotides polymerized during a single binding event or cycle and describes the competition between elongation in the bound state and dissociation from template-primer (75). When DNA polymerases conduct one cycle of synthesis (processive DNA synthesis), the template position for termination of DNA synthesis (pausing) is non-random on heteropolymeric templates (76). Pausing occurs preferentially at certain nucleotide positions and termination probabilities vary at different template positions (77). HIV-1 RT has to polymerize approximately 20,000 nucleotides to complete the conversion of single-stranded RNA genome into double-stranded proviral DNA (78). Following infection by HIV in culture, synthesis of the proviral DNA is completed within 4 to 8 hours (79, 80). Thus, HIV-1 RT polymerizes DNA at an average rate of about 1 nucleotide per second in vivo. In vitro, the processivity of HIV-1 RT ranges from about 1 nucleotide on poly dA to about 300 nucleotides on poly rA and on some M13 templates (67, 81). HIV-1 RT processivity is relatively low compared to the other replicative

polymerases and is observed to be strongly influenced by the template sequences (82). Studies of the structure of HIV-1 RT and its relationship with the processivity and pausing during DNA synthesis are important for understanding viral replication in vivo. These studies should include the evaluation of the ability of template sequence to influence termination of DNA synthesis by HIV-1 RT.

Contribution of RNase H domain of p66 to processivity of DNA synthesis by HIV-1 RT

As described previously, the RNase H domain of HIV-1 RT resides in the carboxyl-terminal part of the protein. HIV-1 RT RNase H function has been identified as involving endo- as well as exonuclease activity (63). The carboxyl terminal region of HIV-1 RT has been recovered from an enzyme preparation from virus particles as an indepedent p15 protein, and a combination of recombinant p15 and p51 leads to reconstitution of RNase H activity (83, 84). The coordination between RNase H activity and RNA-dependent DNA synthesis has been investigated. Models in which there is a tight coupling between DNA polymerization and RNA hydrolysis have been proposed (64, 85-86). However, evidence that RNase H function is less active than the polymerase function during processive DNA synthesis supports the uncoupled model (87, 88). Defects in RNase H function were found to disturb DNA polymerase activity of HIV-1 RT. A Gln475Glu mutant of HIV-1 RT, which has partially impaired RNase H activity, failed to hydrolyze homopolymeric substrates and showed a reduced processivity of the RNA dependent DNA polymerase (89). Two mutants of HIV-1 RT, Asp443Asn and Asp443Asn/Asp498Asn lost detectable RNase H activity but appeared to retain wild-type activity of DNA synthesis on homopolymeric RNA templates (90). However, on heteropolymeric RNA templates derived from the *gag* region of HIV-1, these mutants could poorly extend the terminated nascent

primer. It has been suggested that a catalytically functional RNase H domain is required to facilitate the reinitiation of DNA synthesis at specific pause sites along a heteropolymeric template (90). This suggestion is supported by the observation that the inability of the RNase H negative mutant to hydrolyze the RNA template blocked further DNA synthesis (91). It has also been suggested that conformation alteration induced by mutations in the RNase H domain can affect the favorable primer-template binding required for processive DNA synthesis (91). Bavand *et al.* have studied the polymerization properties of the p51 homodimer. They concluded that the principal effect of the RNase H domain on polymerase function is that it contributes to the stabilization of the p51 subunit conformation in the heterodimer which is necessary for formation of the Michaelis complex (92).

Contribution of P51 to HIV-1 RT activity

Sequence alignments of RTs and other polymerases have suggested that the p66 subunit of the HIV-1 RT contains all the sequence information needed for activity (48). Also, it has been shown that the p66 chain provides the polymerase active site in the p66/p51 heterodimer (41,49). The monomeric forms of the HIV-1 RT, p66 and p51, have been shown to be enzymatically inert (93). Dissociation of either heterodimeric (p66/p51) or homodimeric (p66/p66 and p51/p51) recombinant HIV-1 RT leads to loss of RNA-dependent DNA polymerase activity in the resulting monomer, and it was reported that homologous HIV-1 RT subunits can rapidly associate to form active heterodimer (94). Hence monomeric p66, although it contains all the sequence information needed for activity, requires the interaction with a second polypeptide to maintain the favorable conformation for enzyme activity.

In the crystal structure of HIV-1 RT, the folding of the p51 polypeptide was found to be astonishingly different from that of p66 (50). Unlike p66, p51 has no cleft for the primer-template binding, and the residues involved in catalysis are essentially buried. The two subunits interact in a manner such that the tip of the fingers and the connection subdomains of p51 interact with the palm and connection subdomains of p66 while the very extended thumb of p51 contacts the RNase H domain of p66. It is conceivable that the p66 monomer by itself would adopt a conformation similar to p51 and that this conformation might lack the physical cleft for primer-template binding. However, the protein-protein interactions of p66 with p51 (or p66 in homodimer) enables p66 to induce an enzymatically active conformation.

Effect of primer length on affinity for substrate

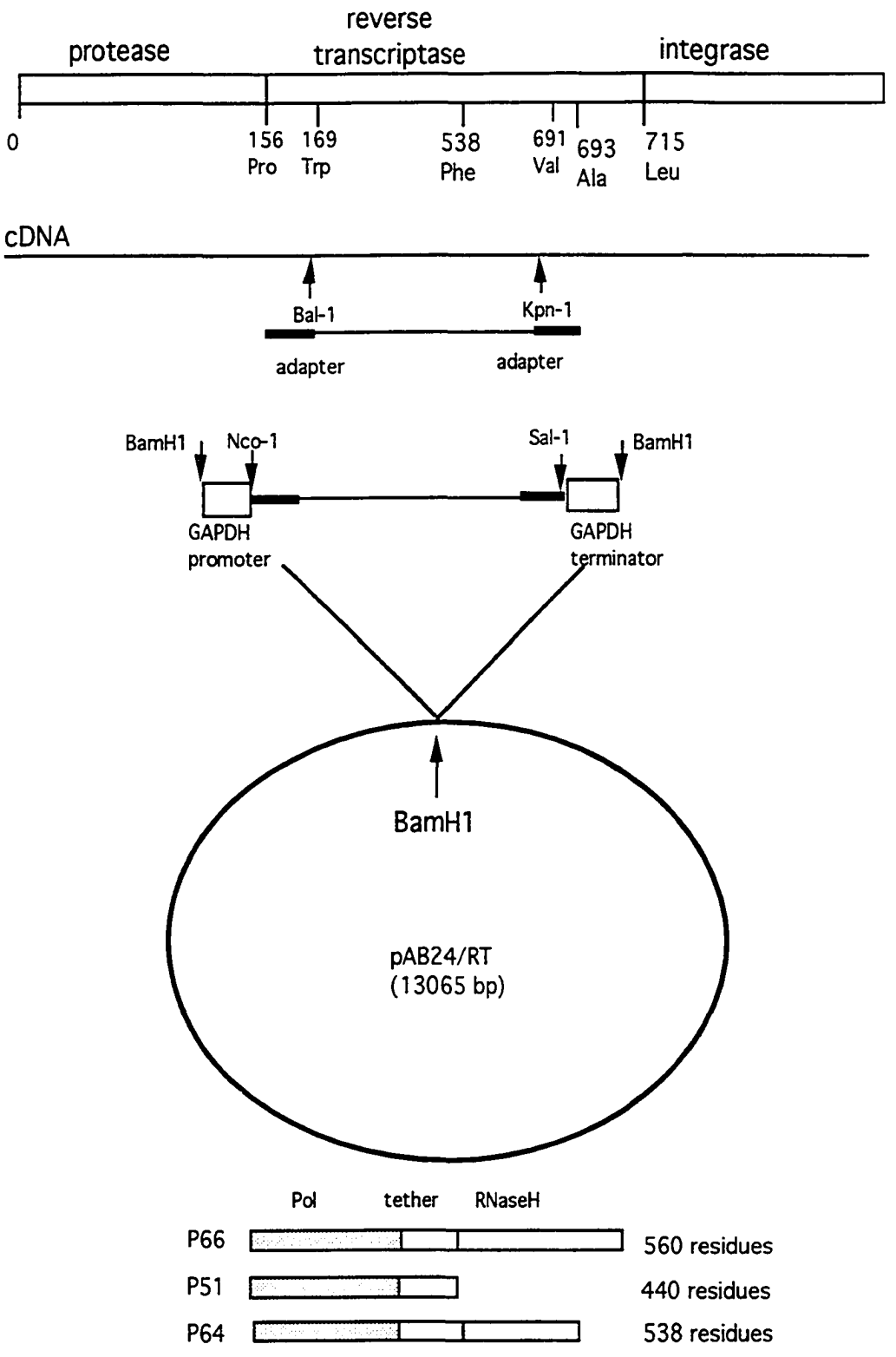
It has been shown that the K_d for poly(rA)$(dT)_n$ is dependent on the length of the primer, with short primers binding with a lower affinity (74). Short oligo (dT) primers (less than 14 nucleotides), annealed to poly(rA), bound RT relatively weakly (K_d = 20 - 30 nM), whereas oligo (dT) primers of 16 and 20 nucleotides in length bound tightly to RT (K_d ~ 0.1 nM) (74). The weak affinity is thought to be due to the fact that the dissociation of the RT-template-primer complex is much more rapid when the primer is short (74). Since, from the computer substrate binding model, the RNA-DNA duplex needs to contact both DNA polymerase and RNase H active sites and fit into a groove mainly folded by the p66 subunit, the short primer might cause insufficient interaction with the substrate binding groove and consequently make the primer-template easier to dissociate from RT than the substrate with long primer.

Significance of this investigation

The modeling of the binding of an RNA/DNA strand into the cleft shown in the crystal structure (50) suggested that significant contacts between the RNase H domain and the primer-template are occurring. As described before, point mutations in the RNase H domain can impair the RNase H function and hinder DNA synthesis as well (89-91). It is important to investigate the correlation between primer-template binding and DNA synthesis of the RNase H negative enzyme.

In order to study the structure and function relationships of different forms of HIV-1 RT, a well characterized recombinant C-terminal deletion mutant of HIV-1 RT will be employed. This mutant enzyme of HIV-1 RT has been cloned and expressed in S. cerevisiae (35,95). A brief description of the genetic manipulation of this mutant clone as well as the p66 and p51 is given in Figure 6. A DNA fragment from restriction digestion at sites Bal-1 (Trp179) and Kpn-1 (Trp690) of the HIV *pol* open reading frame was fused with synthetic oligonucleotide adapters for desired coding regions. The N-terminus is expanded to Pro156 and the C-terminus to Ala693, Leu715, and Phe595 to construct three clones: p64, p66, and p51 respectively. The RT gene is flanked by promoter and terminator sequences of the yeast glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene and inserted into pAB24, an autonomously replicating yeast plasmid containing LEU2 and URA3 markers for selection. The purified polypeptide, p64, which is missing of 22 amino acid residues in the RNase H domain of p66, lacks RNase H activity but retains the DNA polymerase activity (35, 95, 96). Heard *et al.* have demonstrated that p64 has the standard ordered Bi-Bi- mechanism as shown in wild-type enzyme with less affinity for the primer-template (97).

Figure 6 Schematic relationship of HIV-1 RT clones (35,95).



In this investigation, the specific aims are as follows :

1. *Reconstitution of the different forms of HIV-1 RT.*

P66, p51, and p64 (the C-terminal deletion mutant) will be used to generate wildtype (p66/p51) heterodimer, p66/p66 homodimer, p64/p64 homodimer, and p64/p51 heterodimer by a reconstitution method. The interactions of p66 or p64 with p51 as heterodimers (p66/p51 and p64/p51) will be examined. The increase in DNA polymerase activity in both wildtype and mutant heterodimers of HIV-1 RT relative to the homodimers will be expected if the catalytic subunits are interacting with p51 such that a favorable structure for DNA synthesis is adopted.

2. *Investigation of the function of the C-terminal region in p66 on primer-template binding.*

The dissociation constant for primer-template will be studied among these four forms of HIV-1 RT. p66/p66 is the form of HIV-1 RT whose small subunit contains an extra stretch of RNase H domain (p15) instead of p51 only as in the wildtype. The extra p15 might cause p66, the catalytic subunit, to adopt a conformation which affects primer-template binding. Also, the extra p15 might have a physical contact with primer-template and hence influence the primer-template binding. By comparing homodimers with p66/p51 heterodimers, the effect of this extra p15 on primer-template binding will be examined. P64/p51 is the form of HIV-1 RT which has 22 amino acid residues deleted at the C-terminus of p66. The truncated RNase H domain might cause the loss of the affinity for primer-template. This dissociation constant for primer-template of this mutant heterodimer will be compared to that of wild-type enzyme. p64/p64 is the form of HIV-1 RT which has not only an extra stretch of polypeptide on the small subunit, but also a 22 amino acid residue deletion in the RNase H domain of p66. The effect of these double defects on the primer-template binding will be examined.

3. *Investigation of the function of C-terminal region in p66 on processivity.*

The processivity of RNA-dependent DNA synthesis will be compared among these different forms of HIV-1 RT. The extra stretch of polypeptide in p66/p66 might have altered the conformation of the catalytic subunit and hence influence the processivity of the enzyme. In p64/p51, the structure of the small subunit is as that in the wild-type, but the deletion in the catalytic subunit causes a defect in RNase H function. p64/p51 is not able to hydrolyze the RNA on the RNA-DNA duplex and hence might create blocks for DNA synthesis and hinder processivity. In p64/p64, the catalytic subunit assumes an altered conformation due to the interaction with p64 and lacks RNase H function. The processivity of this mutant enzyme will be compared to those in wild-type enzyme and p64/p51. Besides using poly(rA)(dT) as the primer-template for this investigation, a random base RNA template is synthesized to serve as a template for DNA synthesis by HIV-1 RT. The significance of this template is that it is a natural substrate as that in vivo and, since the sequence of this template is known, the relationship between the primer extension pattern and the template sequence can be investigated.

4. *Investigation of the function of the C-terminal deletion in p66 on primer-extension pattern.*

The relationship between the processivity of DNA synthesis by HIV-1 RT and RNA template sequence will be investigated. The synthesized RNA template is 335 bases long and has a random sequence of ribonucleotides. By hybridizing the RNA template with a DNA 20-mer, 78 nucleotides can be incorporated to the 3' end of the primer. The primer extension pattern by p66/p66, p66/p51, p64/p64 and p64/p51 will be examined. The template sequence where DNA synthesis pauses will be compared.

5. *Investigation of the effect of primer length on the primer-template binding to p64/p51 mutant heterodimer*

As described before, Beard *et al.* have reported that the K_d for poly(rA)$(dT)_n$ is dependent on the length of the primer with lower affinity for the short primer (74). The primer-template binding of p64/p51 will be characterized for its dependence on the primer length. A DNA 12-mer will be annealed to the RNA335 template. The K_d for RNA335-DNA12 and RNA335-DNA20 will be compared to investigate if the C-terminal deletion mutant shows lower affinity for the primer-template with shorter primer as in the wild type.

II. Materials and Methods

Materials

Adenine, Tyrosine, Lysine, Uridine, Phenylalanine, Tryptophan, Methionine, Arginine, and **Histidine**: Purchased from Sigama.

AMV-RT (8,000 U/ml): Purchased from Promega.

Bacto- agar: Purchased from DIFCO.

β-mercaptoethanol: Purchased from BioRad.

BSA: Purchased from Promega as nuclease free stock solution at 10 mg/ml.

Coomasie plus protein assay reagent: Purchased from Pierce.

CytoScint scintillation cocktail: Purchased from ICN.

2'-Deoxythymidine-5'-Triphosphate (dTTP),

2'-Deoxyadenosine-5'-Triphosphate (dATP),

2'Deoxyguanosine-5'-Triphosphate (dGTP),

and **2'-Deoxycytidine-5'-Triphosphate (dCTP)**

: Purchased from Sigma and prepared as a 100 mM stock solution .

Dithiothreitol (DTT): Purchased from Fisher are prepared as a 1.0 M stock solution.

DNA 5'-End Labeling System: Purchased from Promega.

[gamma-^{32}P] Adenosine-5'-Triphosphate: Purchased from ICN Biomedicals, INC., 10 mCi/ml, 2.2 μM.

Gel filtration standards: Purchased from BioRad.

GF/C glass microfibre filters: Purchased from Fisher.

Guanosine-5'-diphosphate: Purchased from Boehringer Mannheim.

Guanyl (3'-5') uridine: Purchased from Sigma.

KCl: Purchased from Fisher and prepared as a 4 M stock solution.

LB broth: Purchased from DIFCO.

[Methyl,1',2',-^{3}H]-Thymidine-5'-triphosphate ([^{3}H]-dTTP): Purchased from Amersham,1.0 mCi, 8 µM, 1.0 mCi/ml.

$MgCl_2$: Purchased from Fisher and prepared as 1 1.0 M stock solution.

NucTrap Push Columns: Purchased from Stratagene.

Oligomers (5' GCT GGC GAA GGG GGG ATG TG 3', 5' CTG GCA CGA CAG GTT TCC CG 3' for PCR, 5' GGT CCA CGG TAT CGA TAA GC 3' for RT assays): Purchased from The Midland Certified Reagent Company.

Peptone: Purchased from Sigma.

Pfu: Purchased from Stratagene.

Phenol (Molecular Biology Grade): Purchased from Fisher.

Phenylmethylsulfonyl fluoride (PMSF): Purchased from Sigma.

Polynucleotide phosphorylase: Purchased from United States Biochemical, 3.2 U/mg.

Polyriboadenosine-5'-monophosphate-Oligothmidine-5'monophosphate poly(rA)$(dT)_{12-18}$: Purchased from Pharmacia, $S_{20,w\ for\ poly(rA)}$= 5.8, poly(rA) to $(dT)_{12-18}$ was 1:1.

Pvu II: Purchased from Promega.

Ribomax Large Scale RNA Production System -T7: Purchased from Promega.

RNA molecular weight marker: Purchased from Boehringer Mannheim Biochemica.

Sodium Pyrophosphate : Purchased from Fisher

S&S membrane filters: Purchased from Schleicher & Schuell, pore size : 0.45 µm, size : 24 MM.

Standard dialysis tubing (MWCO = 14,000): Purchased from Fisher.

Sterile filter: Purchased from Nalgene.

Superose 12 column: Purchased from Pharmacia.

100%Trichloroacetic acid (TCA): Purchased from Fisher in 500 g lots. Stock solution was prepared by the addition of 227 ml of MilliQ water to the bottle.

Tris Base: Purchased from Fisher and prepared as a 1 M stock solution at desired PH.

Triton X-100: Purchased from Fisher

Tryptone: Purchased from DIFCO.

TSK-Gel G3000SW column: Purchased from TosoHaas.

Ultrafiltration membrane: Purchased from Amicon.

Uridine-5'-diphosphate: Purchased from United States Biochemical.

Yeast extract: Purchased from DIFCO.

Yeast Nitrogen base w/o Amino Acids: Purchased from DIFCO.

Methods

Preparation of Leu (-) supplement

The Leu (-) supplement was prepared by combining 0.8 g of Adenine, 0.6 g of Tyrosine, 0.6 g of Lysine, 0.6 g of uridine, 1 g of Phenylalanine, 0.4 g of Tryptophan, 0.4 g of Methionine, 0.4 g of Arginine, and 0.4 g of Histidine.

Preparation of Leu (-) rich media

The Leu (-) rich media was prepared by combining 10 ml of 10X yeast nitrogen base w/o amino acids (prepared by dissolving 6.7 g of nitrogen base in 100 ml of MilliQ water), 0.5 g of Leu (-) amino acid supplement, 5 ml of 1%

Tryptophan, 25 ml of 0.2% Adenine, 400 ml of 20% glucose, and 8 ml of 5% Threonine. The total volume was brought up to 1 liter with MilliQ water. The solution was sterilized by filtering through sterilizing membrane.

Preparation of YEP media

The YEP media was prepared by combining 20 g of peptone, 10 g of yeast extract, and 920 ml of MilliQ water and autoclaving the solution. 25 ml of 0.2% Adenine, 5 ml of 1% Tryptophan and 50 ml of 20% glucose were then added to the solution.

Yeast cell culture

The yeast cell culture was prepared by inoculating 20 ml Leu(-) rich media with 2 ml of yeast cells in a 250 ml bottle and shaking the culture at 30°C for 24 hours. The culture was transferred to a 1 liter bottle containing 100 ml of Leu (-) rich media and shaken at 30°C for 24 hours. The culture was further transferred to a 2 liter flask containing 1 liter of YEP media and shaken at 30°C for 30-36 hours. The yeast cells were then pelleted by centrifugation at 10,000 rpm in JA-14 rotor at 4°C for 20 minutes. The cell pellets were stored at - 70°C.

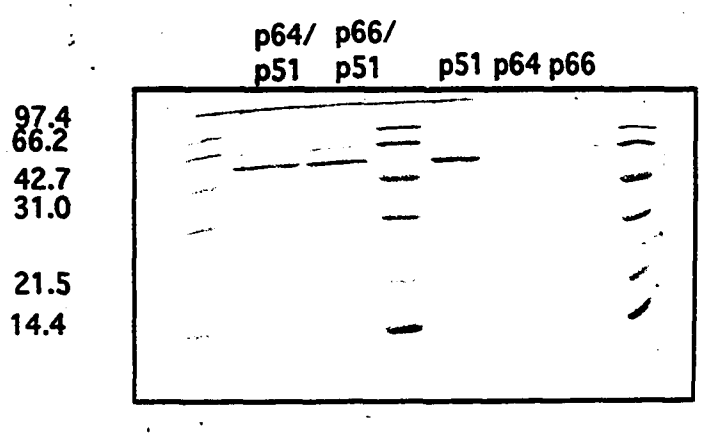
Preparation of yeast cell lysis buffer

The yeast cell lysis buffer was prepared by combining the following quantities of stock solutions : 100 ml of 10X RT buffer (500 mM Tris, pH 7.8 , 2 mM EDTA, 1% Triton X-100, and 500 M KCl), 700 ml of MilliQ water, 200 ml of glycerol, 3.7 g of KCl, 1.41 g of EDTA, and 154 µl of β-mercaptoethanol. The final solution contains 50 mM Tris, pH 7.8, 0.1% Triton X-100, 100 mM KCl, 20% glycerol, 4 mM EDTA, and 10 mM β-mercaptoethanol.

Protein purification

The cloning of HIV-1 RT: pAB24/RT5 (p66), pAB24/RT6 (p51), and pAB24/RT4 (p64), in yeast has been described previously (Figure 6). The pellets from yeast cell culture were used for RT purification. Yeast cell pellets (from 1 liter yeast cell cultures) were broken by glass bead lysis in 200 ml of lysis buffer in the presence of 1 mM PMSF. The lysis mixture was centrifuged at 10,000 rpm, 0°C for 30 minutes. The supernatant was transferred to a cold beaker and $(NH_4)_2SO_4$ (39.8 g per 100 ml of the supernatant) was added to salt out the proteins. The mixture was pelleted by centrifugation at 10,000 rpm, 0°C for 30 minutes. The muddy pellets were then transferred into a dialysis tubing (MWCO = 14,000) and dialyzed in 4 liters of 1X RT buffer (50 mM Tris, pH 7.8, 0.2 mM EDTA, 0.1% Triton X-100, 50 mM KCl, 10% glycerol, 10 mM β-mercaptoethanol) overnight. The desalted mixture was then applied to a Whatman P11 phosphocellulose column equilibrated in 1X RT buffer and eluted with a linear 50-800 mM KCl gradient. Fractions containing RT were determined by SDS-PAGE (10%), protein assay, and activity assay, and these fractions were pooled. Pooled fractions were concentrated and desalted in an Amicon pressure cell with a YM30 membrane (Amicon, MWCO = 30,000 MW) under 50 psi at 4°C. The concentrated fractions were then applied to an Affi-Blue column equilibrated in 1X RT buffer without KCl. The column was washed with one bed volume of RT buffer (no KCl). Proteins were eluted with a linear 50-600 mM KCl gradient. Active fractions were pooled and concentrated as before. If necessary, RT was further purified by gel filtration on a Superose 12 column (Pharmacia). The elution buffer contains 100 mM Tris, pH 8.5, 10 mM β-mercaptoethanol, 1 mM EDTA, 0.1% Triton X-100, 50 mM KCl, and 10% glycerol. The purified protein was concentrated to 1-2 μg/μl, divided into 50 μl aliquots and stored frozen at -70°C. The purified enzyme was analyzed on SDS-PAGE as shown in Figure 7.

Figure 7 SDS-PAGE analysis of p66, p64, p51, p66/p51, and p64/p51. 1 μg of each RT homodimer and 2 μg of each RT heterodimer were analyzed on an SDS polyacrylamide gel (10%) stained with Coomassie blue. Molecular weight standards shown on the gel are 97.4, 66.2, 42.7, 31, 21.5, and 14.4 KD.



Also, the relationship between the amount of dTMPs incorporated into poly(rA)(dT) template-primer and enzyme concentration was shown in Figures 8 and 9 for p66/p66 and p64/p64 respectively to determine the optimal RT dilution for further studies.

Protein assay

The protein assay used for determining the protein concentration during the RT purification is based on the Bradford method which utilize an absorbance shift in an acidic Coomassie Brilliant Blue G-250 solution. An absorbance shift from 465 to 595 nM occurs when Coomassie Blue G-250 binds to proteins in an acidic solution. The color response is relatively linear to the protein concentration. 0-10 μg/ml of BSA was used to obtain the linear relationship between the absorbance at 595 nM and protein concentration. The concentration of RT sample was obtained by the absorbance and the linear equation of the standard curve.

Preparation of LB (Luria - Bertaini) medium

The LB medium contains 10 g of Bacto-tryptone, 5 g of Bacto-yeast, and 10 g of NaCl in a water solution with final volume of 1 liter. The medium was autoclaved. Ampicillin was added to 100 μg/ml after autoclaved solution was cooled to 50°C.

Preparation of LB plate

15 g of agar was added to 1 liter of LB medium and autoclaved. Ampicillin was added to 100 μg/ml when the solution was cooled to 55°C for LB-Amp plate.

Figure 8 dTMP incorporation by purified p66. Poly(rA)$(dT)_{12-18}$ was used as the template-primer and different dilutions of purified p66 were used in the assays.

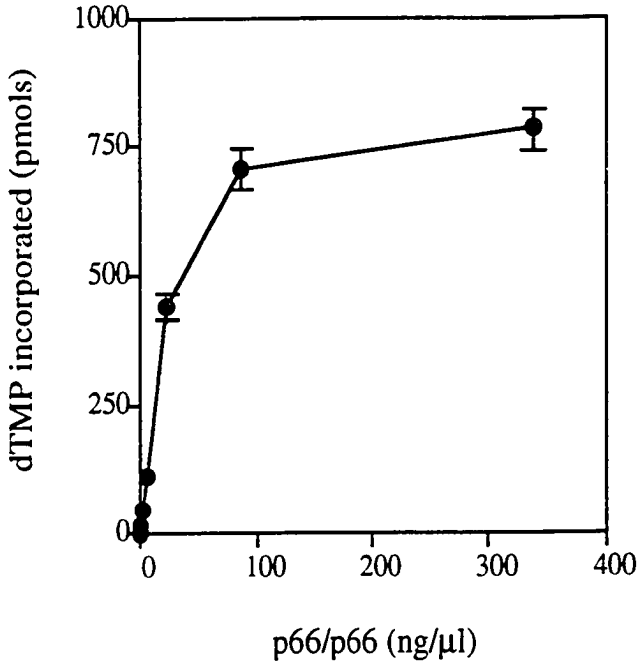
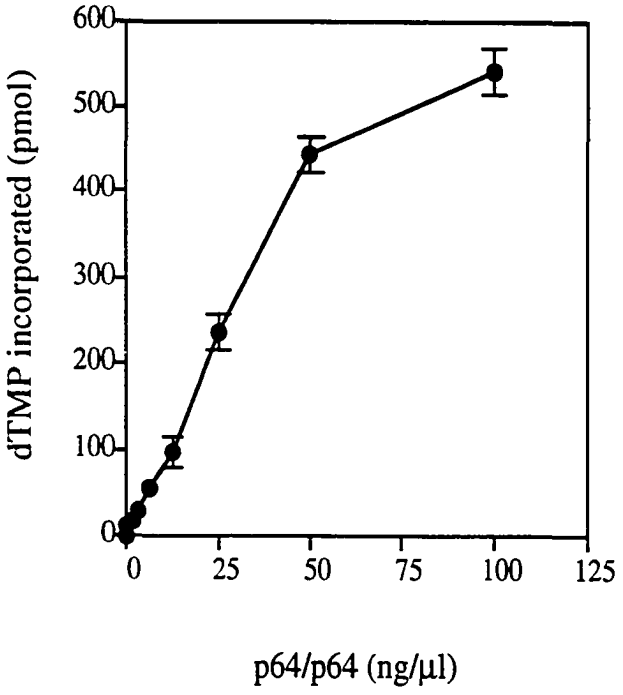


Figure 9 dTMP incorporation by purified p64. Poly(rA)$(dT)_{12\text{-}18}$ was used as the template-primer and different dilutions of the purified p64 were used in the assays.



30-35 ml of medium was poured into each 100 mm petri dish. After the plates hardened, they were stored at 4°C .

Preparation of HB101 competent cells

1 ml of bacteria HB101 was inoculated into 10 ml of LB medium and incubated overnight at 37°C in a shaking chamber. The culture was streaked on an LB plate and incubated overnight at 37°C. One bacterial colony from the LB plate was inoculated into 5 ml LB medium and incubated overnight at 37°C in a shaking chamber. 1 ml of the bacterial culture was used to inoculate 100 ml of LB incubated at 37°C in a shaking chamber for 2-4 hours until OD_{550} = 0.5. 50 ml of the culture was put into each of two sterile centrifuge tubes, and they were chilled on ice for 10 minutes. The cells were pelleted by centrifugation at 4,000 x g (7,000 rpm) at 4°C for 5 minutes. They were resuspended in 25 ml of ice-cold, sterile trituration buffer (50 mM $CaCl_2$ and 10 mM Tris, pH 8.0). Placed in an ice bath for 15 minutes, and centrifuged as before. The supernatant was decanted and cells resuspended in 3.3 ml of trituration buffer. The final suspension of cells was stored at -70°C.

Transformation of HB101 with plasmid Bluescript KS(+)

A 200 µl aliquot of HB101 competent cells was thawed on ice and 1-6 µl of plasmid (40 ng/ml) was added. Cells were left on ice for 30 minutes and then heat shocked at 42°C for 2 minutes. 1 ml of LB was added to each tube and the mixtures were incubated at 37°C for 1 hour without shaking. 50 µl of the culture was spread on LB-Amp plate and incubated at 37°C for 16-24 hours.

Isolation of plasmid DNA from HB101/pBS

A 2.5 ml overnight culture from a single colony was used to inoculate a 2-liter flask containing 500 ml of LB-Amp medium. The 500 ml culture was incubated 12-16 hours at 37°C with shaking. The bacterial cells were harvested by centrifugation at 4000 x g (10,000 rpm) at 4°C for 20 minutes. The pellet was recovered and stored at -70°C if not used immediately. The pellet from the 500 ml culture was resuspended in 5 ml of solution I (50 mM glucose, 10 mM EDTA, and 25 mM TrisCl, pH 8.0). After 10 minutes at room temperature, 10 ml of fresh solution II (0.2 N NaOH and 1% SDS) was added. The suspension was stirred and allowed to sit on ice for 10 minutes. 7.5 ml of solution III (3 M potassium acetate and 11.5% glacial acetic acid) was added, mixed, and the solution was chilled for another 10 minutes on ice. After centrifugation at 20,000 x g (13,000 rpm, JA20) at 4°C for 10 minutes, the supernatant was transferred to a fresh tube, 0.6 volumes of isopropanol was added, and the sample was centrifuged again at 15,000 x g (11,500 rpm, JA20). The pellet was washed with 2 volumes of 70% ethanol and dried under vacuum. After resuspension in 4 ml of TE (10 mM TrisCl, pH 7.6, 1 mM EDTA), 4.4 g of CsCl and 0.4 ml of 10 mg/ml Ethidium Bromide were added. The mixture was centrifuged at 350,000 x g (58,000 rpm, VTi 65) for 20 hours. The lower band was collected by a needle and syringe, and the Ethidium Bromide was removed by repeated extraction with isopropyl alcohol. The CsCl was removed by dialysis in TE buffer.

Preparation of 10X RT assay buffer

The 10X RT assay buffer was prepared by combining the following quantities of stock solutions: 1250 µl of 1 M Tris (pH 8.5), 145 µl of 1 M $MgCl_2$, 500 µl of 4.0 M KCl, 200 µl of 1M DTT, 125 µl of 10% Triton X-100, and 280 µl of

MilliQ water. The final volume was 2500 µl. The buffer contains 500 mM Tris, 60 mM $MgCl_2$, 800 mM KCl, 80 mM DTT, and 0.5% Triton X-100.

Preparation of poly(rA)$(dT)_{12-18}$ stock solution

The lyophilized poly(rA)$(dT)_{12-18}$ (25 units) was dissolved by adding 15 µl of 4 M KCl and 781 µl of (10 mM Tris/ 1mM EDTA) to the bottle to make a final concentration of 2.0 µg/ul. The molarity of the poly(rA)$(dT)_{12-18}$ was obtained from the following equation (119): Log(M) = 2.179 (Log ($S_{20,w}$) + 1.523). A molecular weight of 95968 g/mole was obtained from the $S_{20,w}$ of 5.8 in the certificate of analysis. The average base length corresponding to this sedimentation constant is 258 polymerized adenosine-5'-monophosphate bases. Final stock concentration of the poly(rA)$(dT)_{12-18}$ was 20.8 µM.

RT activity assay

The RT activity assays were carried out in a 50 µl mixture containing 50 mM Tris (pH 8.5), 0.05% Triton X-100,10 mM $MgCl_2$, 60 mM KCl, 8 mM DTT, 0.05% Triton X-100, 6 mM $MgCl_2$, 80 mM KCl, 2 µg of poly(rA)(dT), 80 µM dTTP, and 0.25 mCi/nmol [^{3}H]-dTTP. Reactions were initiated by the addition of 10 µl of diluted enzyme to 40 µl of cocktail solution and incubated at 37°C for 30 minutes. The reactions were quenched with 800 µl of ice cold 10% TCA containing 10 mM PPi and maintained at 0°C until precipitated material was collected on Whatman GF/C filters. Each filter was washed with 8 ml of the cold TCA solution and then rinsed with 1-2 ml ethanol. The dry filters were counted in a 5 ml CytoScint scintillation cocktail in a Scintillation Counter. Activity of RT is indicated by the amount of [^{3}H]-dTTP incorporation in 30 minutes at 37°C.

Nitrocellulose Binding Assays

The labeled primer-template was first prepared by using AMV-1 RT to incorporate [^{3}H]-dTTP onto poly(rA)(dT). The labeled primer-template was incubated with RT (0-10 nM) with RTB in a final volume of 800 µl for 15 minutes at 0°C. The mixture was filtered through nitrocellulose filters to capture the enzyme bound primer-template. The filters were placed in vials containing 5 ml of scintillation cocktail and counted on a Beckman liquid scintillation counter. The K_d was calculated by the Eadie plot method.

Poly(U,G) synthesis

Polynucleotide phosphorylase (0.5 units/ml) was incubated overnight at 37°C in a reaction mixture containing 0.15 M Tris, pH 9.0, 1 mM EDTA, 0.5mM Guanyl (3'-5') uridine, 1mM $MgCl_2$, 2 mM GDP, 2mM UDP, and 1 ml of water. The poly(U,G) mixture was extracted with phenol and chloroform and resuspended in 100 µl of TE buffer after ethanol precipitation. Concentration was determined by optical density at 260 nM. Analysis of 5 µg of the sample by agarose gel electrophoresis showed the average size to be 0.5 kb.

Processivity assays on poly(rA)(dT)

Assays to determine the processivity of RTs were carried out by using poly(U,G) as challengers. RT was incubated with poly(rA)(dT) (40 ng/µl) at 0°C for 10 minutes and the reaction was initiated by adding $MgCl_2$ (6mM), [^{3}H]-dTTP (80 µM, 0.25 mCi/nmol), and poly(U,G) (0.1-0.2 mg/ml). The mixture was incubated at 37°C and aliquots of 50 µl were removed at times over a 30 minutes interval for counting to obtain the curve of time course incorporation. Assays without poly(U,G) were also carried out for the control curve. The time required

for one processive cycle was estimated from the point where the control and poly(U,G) curves separate.

Enzyme Reconstitution

Reconstituted heterodimers, p66/p51 and p64/p51, were made by incubating the corresponding enzyme preparations (p66 + p51 and p64 + p51) at 1:1 by weight in a buffer containing 100 mM Tris, pH 8.5, 10 mM β-mercaptoethanol, 1 mM EDTA, 0.1% Triton X-100, 50 mM KCl, and 10% glycerol at 4°C for more than 24 hours. The effect of reconstitution was determined by comparing the activity of incorporating dTMPs onto primer-template. Each activity assay has a volume of 20 µl and contains dNTPs (50 µM), 1 µg of poly(rA)(dT) or RNA335-DNA20 (0.5 µM), 0.5 µCi of [^{3}H]-dTTP, and RT (0-25ng).

Polymerase Chain Reaction

Plasmid Bluescript (+) (the sequence is shown in Appendix 1) was used as a template in a polymerase chain reaction to amplify a 446 bp fragment that contains the T7 promoter sequence. Two primers have the sequences of 5' GCTGG CGAAG GGGGG ATGTG3' and 5'CTGGC ACGAC AGGTT TCCCG 3'. A 100 µl PCR reaction mixture is comprised of 24 µl of water, 10 µl of 10X PCR buffer, 6 µl of 500 µM dNTPs, 2 µl of each 50 µM primer and 1 µl of 278 ng/µl pBluescript. The 10X PCR buffer contains 60 mM Tris, pH 7.5, 500 mM NaCl, 60 mM $MgCl_2$, and 10 mM DTT. The mixture was initiated by heating at 95°C for 10 minutes followed by 30 amplification cycles of denaturing (95°C for 5 minutes), annealing (68°C for 5 minutes) and DNA synthesis (75°C for 5 minutes) steps. The 446 bp DNA mixture was extracted with phenol and chloroform and resuspended in 20 µl of TE buffer after ethanol precipitation. Concentration was

determined by optical density at 260 nM. Analysis of 5 µg of the sample by agarose gel electrophoresis showed the size to be 0.45 kb.

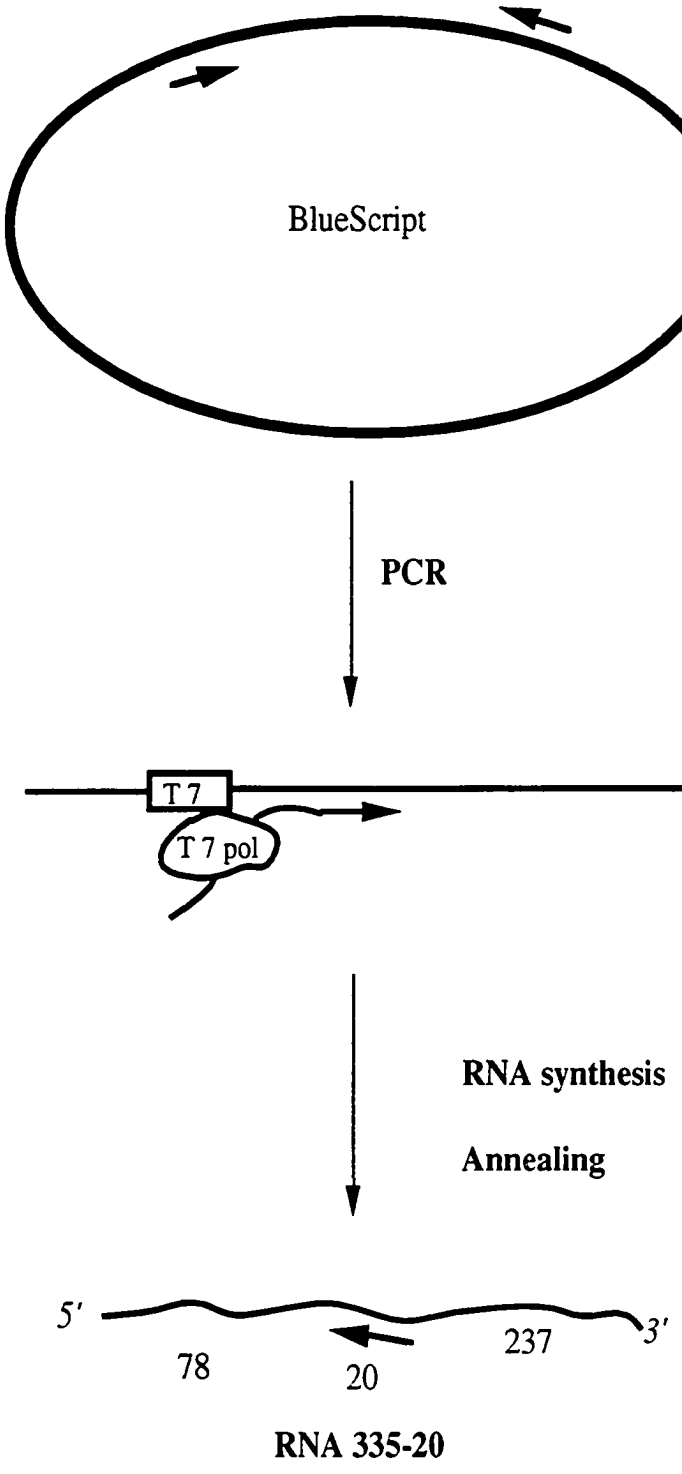
Run-off Transcription

A run-off transcription on the 446 bp PCR product was carried out by using T7 RNA polymerase. The reaction has a total volume of 20 µl and contains 4 µl of 5X transcription buffer, 2.4 µl of 25 mM rNTPs, 2 µl of 0.6 µg/µl DNA template, 10.6 µl of water and 1 µl of 35U/µl T7 RNA polymerase purchased from Promega. The transcription 5X buffer contains 400 mM HEPES-KOH, pH 7.5, 60 mM $MgCl_2$, 10 mM spermidine and 200 mM DTT. The mixture was incubated at 37°C for 4-6 hours. A 335 nucleotide RNA was expected from the reaction. The reaction mixture was first run through a push column to remove unincorporated mononucleotides and then extracted with phenol and chloroform and resuspended in 20 µl of DEPC-treated water after ethanol precipitation. Concentration was determined by optical density at 260 nM. Analysis of 2 µg of the sample by denaturing agarose gel electrophoresis showed the RNA band to fall between the markers of 0.3 and 0.4 kb. The RNA sample was stored at -70°C. The sequence of the RNA produced is shown in Appendix 2. Also, a line diagram and the procedure of generation of this template-primer is shown in Figure 10. The linear relationship between log (kb) and distance migrated is Log (kb) = - 0.411 x distance + 0.719 as shown in Figure 11. The RNA sample has a distance migrated of 2.9 cm and its corresponding size calculated from the linear equation is 0.33659 kb.

5' end labeling of primer

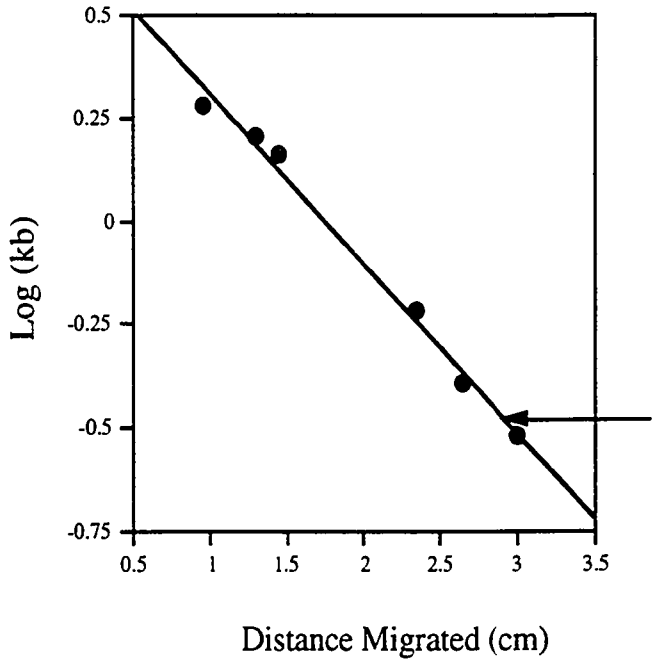
The primers for primer extension reaction were 5' end labeled by [γ-^{32}P]-ATP. In a 10 µl reaction, 10 pmol of primer, 30 µCi of [γ-^{32}P]-ATP and 8 U of T4

Figure 10 A schematic representation of the method used to generate random base template-primer.



RNA 335-20

Figure 11 Plot of size of RNA markers (kb on log scale) against the distance migrated. The arrow indicates the location of the run-off transcription product.



polynucleotide kinase were combined in a 1X buffer containing 50 mM Tris-Cl, pH 9.0, 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$ and 1 mM spermidine. The mixture was incubated at 37°C for 60 minutes and stopped by adding 2 µl of 0.5 M EDTA and 50 µl of TE buffer and heating at 65°C for 5 minutes. The reaction mixture was then run through a push column to remove unincorporated mononucleotides.

Annealing

The RNA335-DNA20 template-primer was prepared by combining 7.5 pmol of DNA20 primer and 2.5 µg RNA335 template in 100 µl of the hybridization buffer containing 50 mM Tris, pH 8.0, 80 mM KCl and 1 mM EDTA. The mixture was incubated at 90°C for 3 minutes and shifted to 60°C to cool down to 25°C in 60 minutes. The annealed RNA335-DNA20 mixture has a final concentration of 75 nM.

Primer extension

Reaction of DNA synthesis on RNA335-DNA20 template-primer by RT was visualized by autoradiography on an 8 M urea, 10% polyacrylamide gel. The primer was 5' end labeled with [γ-^{32}P]-ATP before annealing to the template. Each primer extension reaction contains 10 nM RNA335-DNA20, 50 µM dNTPs and 0.1 µg/µl RT in 1X RT assay buffer. 7.5 µl of the reaction mixture was withdrawn and stopped by the addition of 3 µl of stop solution at 0, 5, 15, 30, 45, 60, 90, 120, 180 and 240 minutes.

Assays to determine K_d for random RNA template primer

The K_d for RNA335-DNA20 and RNA335-DNA12 was obtained by using poly(rA)(dT) as a challenger. 0.1 µg/µl RT enzyme and [^{32}P]-labeled RNA335-DNA20 or RNA335-DNA12 (0 - 45 nM) were incubated on ice for 10 minutes to

form an RT/template-primer complex. 0.2 μg/μl of poly(rA)(dT) and 50 μM dNTPs were used to initiate the DNA synthesis on RT/template-primer complex at 37°C for 30 minutes. In control reactions, poly(rA)(dT) was examined on its ability to trap free RT molecules by combining RT, template-primer, and poly(rA)(dT) at the same time before dNTPs were added to initiate the reaction. 10 μl of each reaction mixture was analyzed on an 8 M urea 10% polyacrylamide gel. The amount of unbound template-primer was evaluated from the radioactivity of the unextended primer. The bands corresponding to the DNA 20-mer or DNA 12-mer were excised, mixed with 5 ml of scintillation fluid and counted. The amount of RT/template-primer complex was calculated by subtracting the amount of RNA335-DNA20 or RNA335-DNA12 used in the reaction by the amount of free template-primer. The Kd was calculated by using Eadie plot method.

III. Results

Reconstitution of p66/p51 and p64/p51

HIV-1 RTs, p66, p64, and p51, were expressed and purified from yeast S. cerevisiae that carries the corresponding gene (Figure 6). According to the previous studies on these three clones, the p66 molecules can associate as homodimers (35) and p66 and p51 were shown to associate as heterodimers (35). Also, it was demonstrated that p66/p66 retained approximately 80% of the RT activity of the heterodimers (35). The C-terminal deletion in the mutant, p64, was shown to abolish RNase H activity (96). But, the mutant retained wildtype DNA polymerase activity and an Ordered Bisubstrate mechanism of polymerization (97). As shown in Figure 6, p66 contains 560 residues: Pro1 to Leu560; p64 contains 538 residues: Pro1 to Ala538; and p51 contains 440 residues: Pro1 to Phe440. In order to study the structure-function relationships of HIV-1 RT, these three clones of RT were reconstituted to the different forms of enzyme : p66/p66, p66/p51, p64/p64, and p64/p51. SDS-PAGE analysis of each individual clone and the reconstituted heterodimer is shown in Figure 7. The effect of heterodimer formation on RNA-dependent DNA polymerase activity was examined by using poly(rA)$(dT)_{12\text{-}18}$ as template-primer. As shown in Figure 12 and Figure 13, the amount of dTMP incorporated into poly(rA)$(dT)_{12\text{-}18}$ increases when the enzyme concentration in the reaction increases and the heterodimer is able to incorporate more dTMP into primer-template relative to homodimer when both are at the same concentration. These results indicate that reconstitution of

Figure 12 Reconstitution of p66/p51 heterodimer increases RT activity. The incorporation of dTMPs on a poly(rA)(dT) template-primer was compared between p66/p66 (■-■) and p66/p51 (●-●) by activity assays. Standard error for each point was less than 5%.

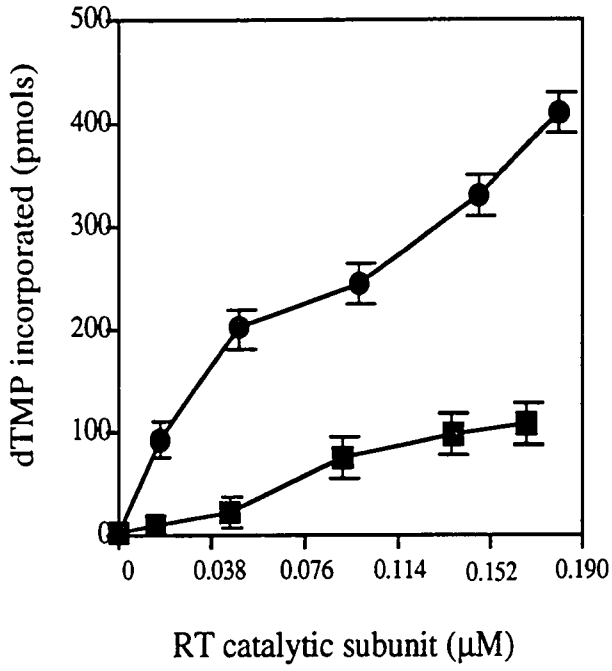
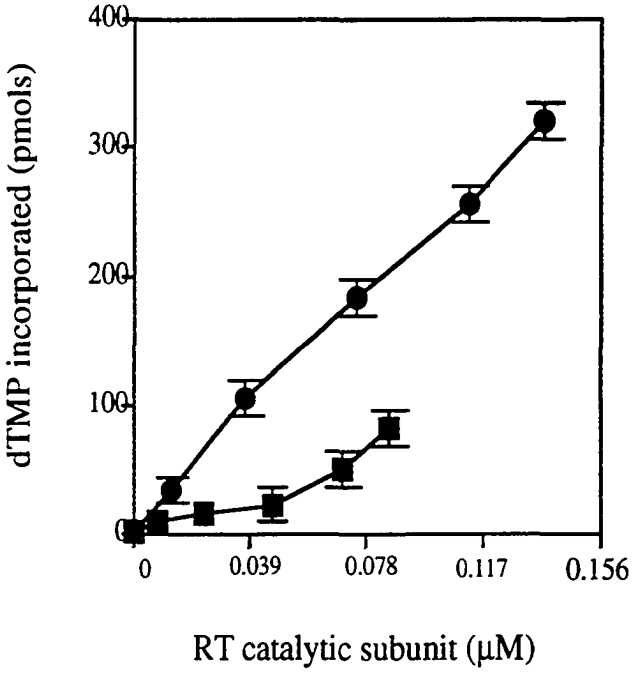


Figure 13 Reconstitution of p64/p51 heterodimer increases RT activity. The incorporation of dTMPs onto poly(rA)(dT) template-primer was compared between p64/p64 (■-■) and p64/p51 (●-●).



p66 or p64 with p51 is able to increase the DNA polymerase activity on poly(rA)$(dT)_{12-18}$. Also, since the RT activity increases when the corresponding subunits were mixed and incubated, subunit interaction in both heterodimer forms, p66/p51 and p64/p51, is stronger than that in either homodimer.

Processivity assays of RTs on poly(rA)(dT)

Processivity refers to the number of nucleotides a polymerase can add to the nascent chain during polymerization before the polymerase dissociates from the primer-template. Poly (U,G) was employed in the processivity assay as a template-primer challenger; an excess of poly (U,G) was added to the assay when the reactions were initiated. Since a preincubation with primer-template preceded reaction initiation, HIV RT has already formed initiation complexes with poly (rA)(dT) and incorporation of [^{3}H]-dTMP will proceed while these complexes are intact. However, when the enzyme dissociates from the new chain on the original primer-template, the excess poly (U,G) "traps" the RT as it binds to a template without a primer or complementary nucleotides. Thus, reinitiation is prevented. The results of these processivity assays are shown in Figures 14-17. As seen in Figure 14, the level of dTMP incorporation by the p66/p66 homodimer continues to increase until approximately eight minutes, at which point the level of incorporation becomes constant in the reaction with challenger. However, the level of incorporation continues to increase normally in the unchallenged control reactions. These results indicate that the homodimer catalyzes nascent chain elongation on the same primer-template until it dissociates after eight minutes; then the free homodimer is trapped by poly (U,G) and unable to incorporate any more dTMP. Similarly, in assays investigating the processivity of the p66/p51heterodimer, p64/p64 homodimer, and p64/p51 heterodimer (Figures 15 - 17), incorporation increases up to fifteen minutes, four minutes, and eight

Figure 14 Assays to determine the processivity of p66/p66 homodimer.
❑-❑ indicates the incorporation of dTMP for unchallenged enzyme assay, and ■-■ indicates the incorporation of dTMP for enzyme assayed in the presence of poly(U,G). The standard deviation of each point is less than 5%.

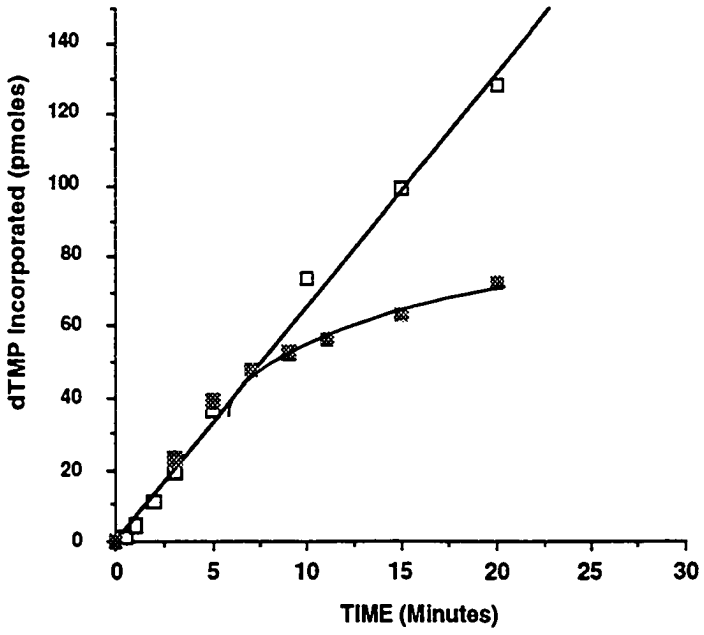


Figure 15 Assays to determine the processivity of p66/p51 heterodimer.
❑-❑ indicates the incorporation of dTMP for unchallenged enzyme assay, and
■-■ indicates the incorporation of dTMP for enzyme assayed in the presence of poly(U,G). The standard deviation of each point is less than 5%.

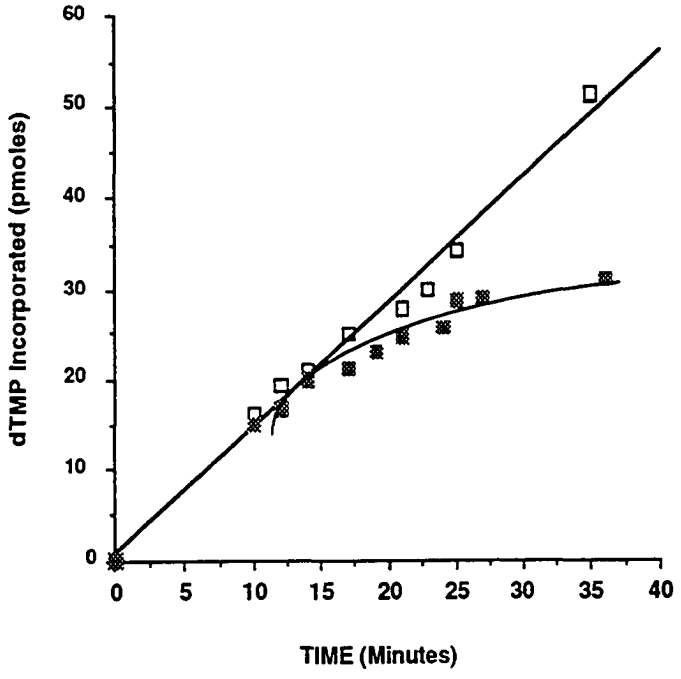


Figure 16 Assays to determine the processivity of p64/p64 homodimer.
○-○ indicates the incorporation of dTMP for unchallenged enzyme assay, and ■-■ indicates the incorporation of dTMP for enzyme assayed in the presence of poly(U,G). The standard deviation of each point is less than 5%.

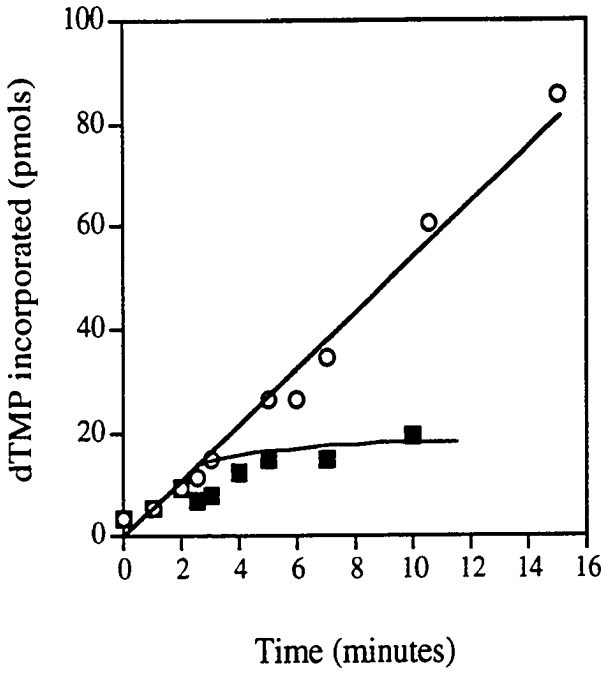
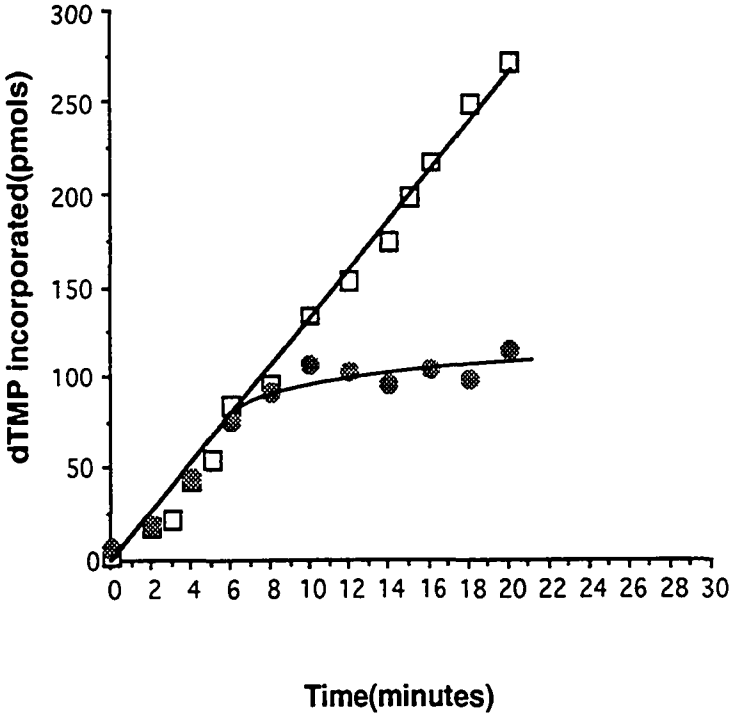


Figure 17 Assays to determine the processivity of p64/p51 heterodimer.
❑-❑ indicates the incorporation of dTMP for unchallenged enzyme assay, and
●-● indicates the incorporation of dTMP for enzyme assayed in the presence of poly(U,G). The standard deviation of each point is less than 5%.



minutes respectively. These observations indicate that the order of processivity of different forms of RT is : p66/p51 (15 min) > p66/p66, p64/p51 (8 min) > p64/p64 (< 4 min). The heterodimer of each wild type and C-terminal deletion mutant was shown to be more processive than the corresponding homodimer.

Nitrocellulose Binding Assays to determine the K_d for Poly(rA)(dT)

The differences in processivity observed among different forms of HIV-1 RT might be related to potentially different affinities of each enzyme form for the primer-template. So, the affinity of different forms of enzyme to poly(rA)(dT) was compared by binding assays. The K_d was calculated based on the Eadie method (100). The binding of RT enzyme (E) to primer-template (L) is decribed by the following equation :

$$E + L \text{ <--> } EL$$

$$Kd = [E][L]/[EL]$$

where [E] and [L] are the concentrations of the unbound enzyme and primer-template, and [EL] is the concentration of enzyme-ligand complex.

In terms of the total enzyme concentration $[E]_0$,

$$[E] = [E]_0 - [EL]$$

the equations

$$[E_0] - [EL] = Kd\ [EL]/[L] \text{ and}$$

$$[EL] = [E_0] - Kd\ [EL]/[L]$$

can be obtained. By plotting $[E_0]$ - [EL] versus [EL]/[L] or [EL] versus [EL]/[L], Kd can be evaluated form the slope of the curve. As shown in Figure 18, the K_d of wild type enzyme for poly(rA)(dT) was obtained by plotting $[E_0]$ - [EL] versus [EL]/[L]. Similarly, as shown in Figure 19, [EL] versus [EL]/[L] was plotted to obtain the K_d for the mutant enzyme. The resulting K_d for p66/p66, p66/p51, p64/p64, and p64/p51 are 16.9$\pm$0.1, 12.8$\pm$0.3, 392$\pm$4, and 21.6$\pm$0.3 nM

Figure 18 Eadie plot of nitrocellulose filter binding assays to determine the affinity of p66/pp66 (▲-▲), p66/p51 (●-●), and p51 (■-■) for poly(rA)(dT). The standard deviation of each point is less than 5%.

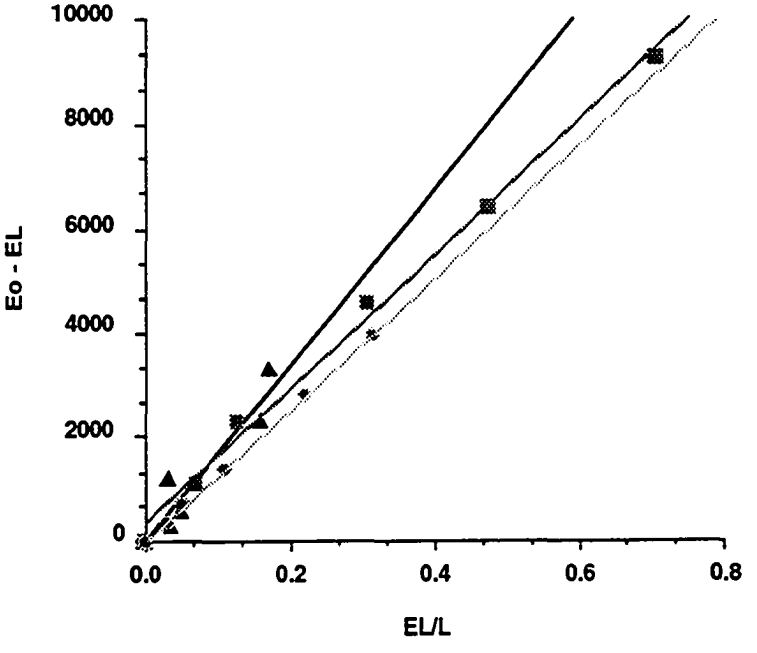
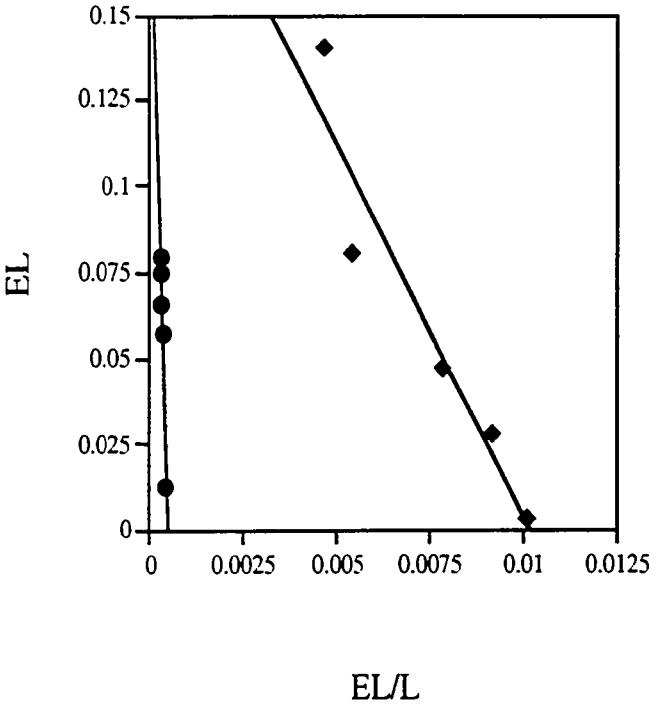


Figure 19 Eadie plot of nitrocellulose filter binding assays to determine the affinity of p64/pp64 (●-●),and p64/p51 (◆-◆)for poly(rA)(dT). The standard deviation of each point is less than 5%.



respectively. The relationship between processivity and affinity for poly(rA)(dT) primer-template is shown in Table 2. Each heterodimer, p66/p66 or p64/p64, has a higher affinity for the primer-template than its homodimer. These results suggest that the increase in processivity seen in the heterodimer is due to the p51 subunit.

Activity assays on RNA335-DNA20 by the reconstituted enzyme

In vivo, HIV-1 RT transcribes the viral RNA genome, which is a random base template (23). Relative to poly(rA), a random base RNA template forms complex secondary structures more easily. The wildtype HIV-1 RT has been shown to have different processivities on random base RNA templates from those on poly(rA) (67,81). Pausing can occur preferentially at certain nucleotide positions during DNA synthesis on a random base RNA template (77). In order to examine the relationship between RT structure and the DNA polymerase function on the random base templates, a 335 nucleotide RNA was generated from plasmid Bluescript by T7 RNA polymerase and used as the template for transcription by RT (Figures 10, 11). A 20 nucleotide oligomer was annealed to this RNA template at a position such that 78 nucleotides could be added to the primer for full length transcription (Figure 10). The incorporation of dNMP was then compared between heterodimer and homodimer to investigate the contribution of the p51 subunit to the enzyme activity on this random base RNA template-primer. As shown in Figures 20 and 21, the results obtained were similar to those in which poly(rA)(dT) was the primer-template. The amount of dNMP incorporated into RNA335-DNA20 increases when the enzyme concentration in the reaction increases and the heterodimer is able to incorporate more dNMP into this primer-template relative to homodimer when both are at the

Table 2 Processivity and Kd of HIV-1 RT for poly(rA)(dT).

Enzyme	p66/p66	p66/p51	p64/p64	p64/p51
Processivity	8'	15'	<4'	8'
Kd (nM)	16.9±0.1	12.86±0.3	392±4	21.6±0.5

Figure 20 Reconstitution of p66/p51 increases RT activity on RNA335-DNA20. The incorporation of dTMPs onto RNA335-DNA20 was compared between p66/p51 (●-●) and p66/p66 (■-■).

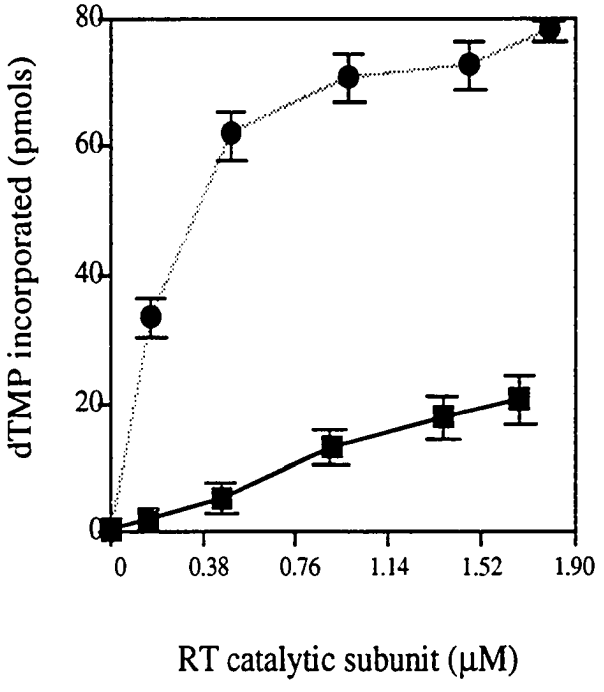
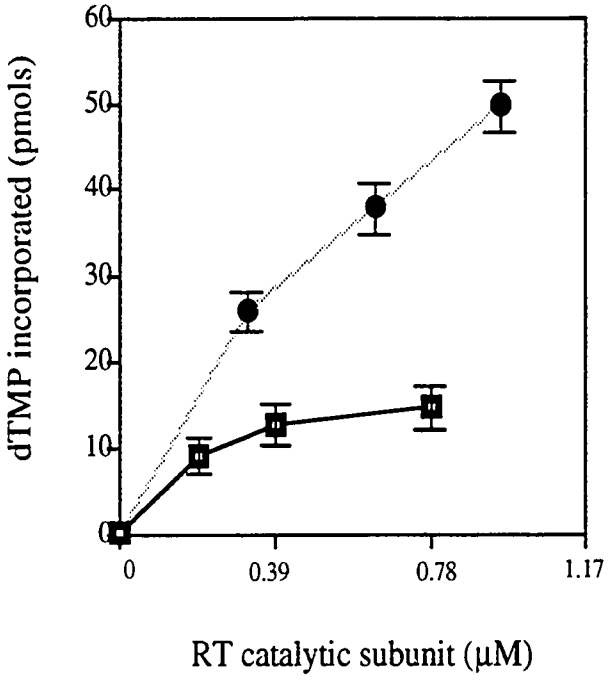


Figure 21 Reconstitution of p64/p51 increases RT activity on RNA335-DNA20. The incorporation of dTMPs onto RNA335-DNA20 was compared between p64/p51 (●-●) and p64/p64 (■-■).



same concentration. These results indicate that reconstitution of p66 or p64 with p51 is able to increase the DNA polymerase activity on RNA335-DNA20.

Primer extension on RNA335-DNA20 by HIV-1 RT

It was observed that heterodimer is more processive than homodimer by using poly(rA)(dT) as the template-primer as described above. It is important to further investigate the processivity on the random base RNA template which is the natural substrate of HIV-1 RT. A random base RNA template can form secondary structure due to hydrogen bonds between bases. Also, the RNA sequence might carry information that will direct the polymerization by HIV-1 RT. The influence of enzyme structure on processivity on a random base RNA template could also be important. In these studies, since the primer was first 5' end labeled with [^{32}P]-ATP before it was annealed to the template, the extension of the primer by RT can be visualized by autoradiography. The labeled RNA335-DNA20 was incubated with RT enzyme for different periods of time so that the effect of reaction time on extension pattern can be observed. DNA size markers were used to indicate the size of the extended primer. Also, by referring to the RNA template sequence, the sequences of pause sites were obtained. The primer extension patterns by homodimer and heterodimer of both wild type and mutant enzyme were compared. The elongation of nucleotides was visualized by an 8 M urea 10% polyacrylamide gel followed by exposure to the X ray film. The results (Figures 22-25) showed that the DNA products extended by p64/p64 homodimer were mainly 38 nucleotides long with a small fraction of the products at 23 to 25 nucleotides long. Unlike p64/p64, however, p66/p66 extended the primer to mostly 38 nucleotides long. Mutant heterodimer, p64/p51, was able to further extend the primer to 62 nucleotides long with small pausing at 43 to 44 nucleotides long. In comparison, the wildtype heterodimer was able to extend

Figure 22 Primer extension by p64/p64 on RNA335-DNA20. Time course of RNA-dependent DNA synthesis on the random base template by RT is shown on an 8 M urea 10% polyacrylamide gel. Reactions were stopped at 0, 5, 15, 30, 45, 60, 90, 120, 180, and 240 minutes as described under Materials and Methods (lanes 2 - 11). The primer (20-mer) alone (lane 1) was used to indicate the position of the unextended primer. Lanes 12 - 14 are DNA size markers which indicate the position of 25, 31, 33, 34, 35, 44, and 61-mer and were used to determine the position of pause sites.

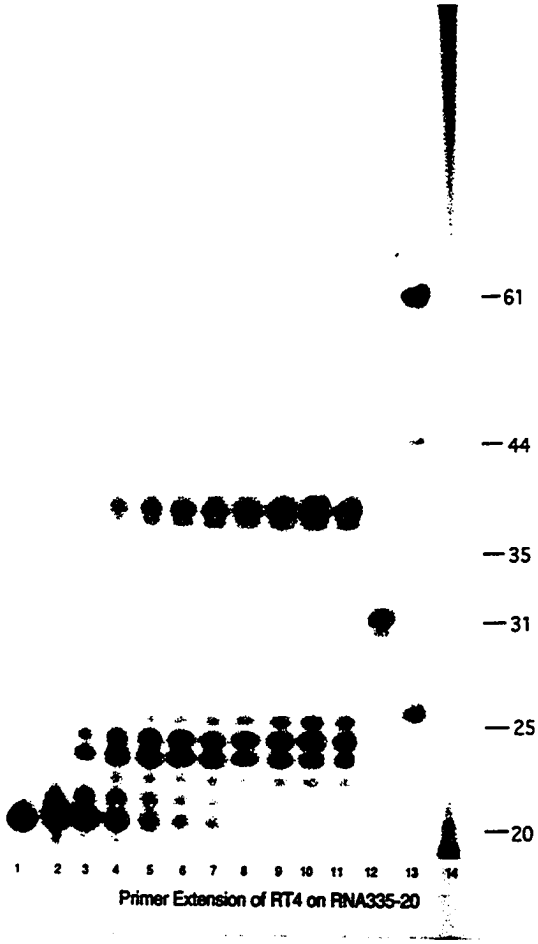
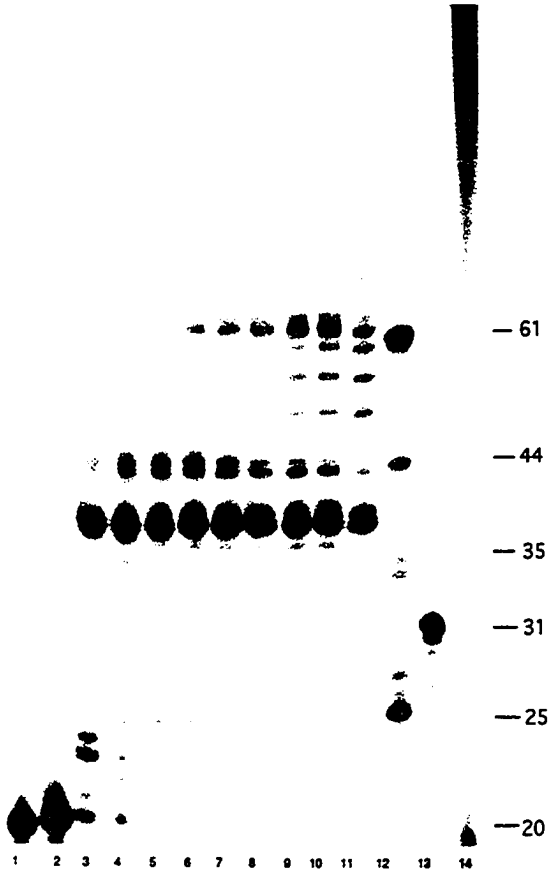
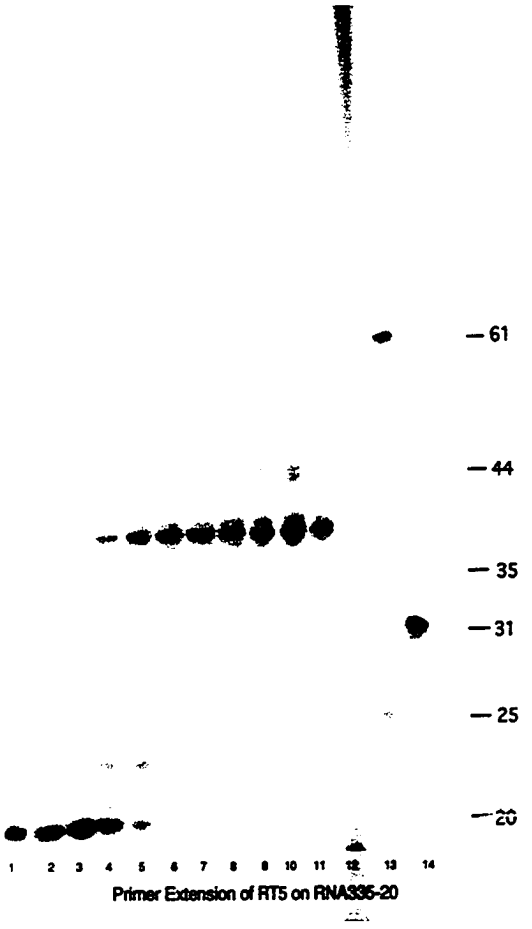


Figure 23 Primer extension by p64/p51 on RNA335-DNA20. Time course of RNA-dependent DNA synthesis on the random base template by RT is shown on an 8 M urea 10% polyacrylamide gel. Reactions were stopped at 0, 5, 15, 30, 45, 60, 90, 120, 180, and 240 minutes as described under Materials and Methods (lanes 2 - 11). The primer (20-mer) alone (lane 1) was used to indicate the position of the unextended primer. Lanes 12 - 14 are DNA size markers which indicate the position of 25, 31, 33, 34, 35, 44, and 61-mer and were used to determine the position of pause sites.



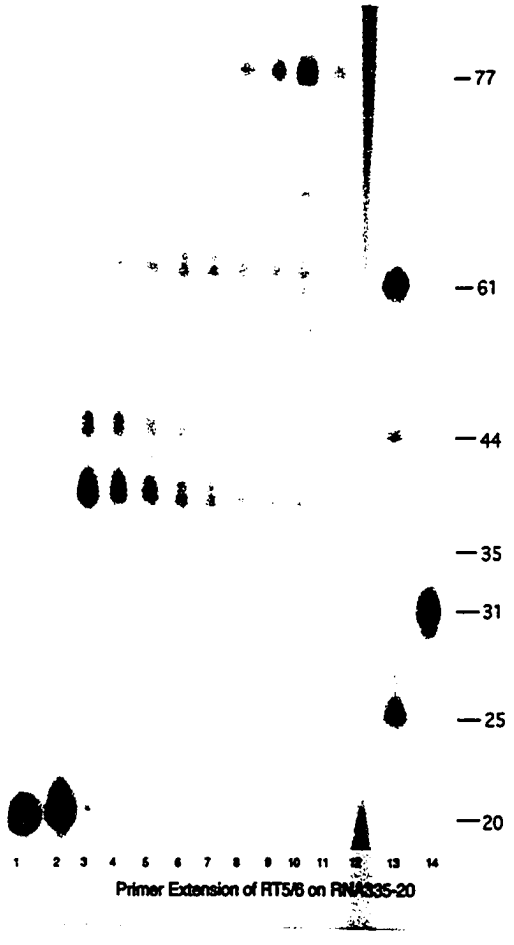
Primer Extension of RT4/6 on RNA335-20

Figure 24 Primer extension by p66/p66 on RNA335-DNA20. Time course of RNA-dependent DNA synthesis on the random base template by RT is shown on an 8 M urea 10% polyacrylamide gel. Reactions were stopped at 0, 5, 15, 30, 45, 60, 90, 120, 180, and 240 minutes as described under Materials and Methods (lanes 2 - 11). The primer (20-mer) alone (lane 1) was used to indicate the position of the unextended primer. Lanes 12 - 14 are DNA size markers which indicate the position of 25, 31, 33, 34, 35, 44, and 61-mer and were used to determine the position of pause sites.



Primer Extension of RT5 on RNA336-20

Figure 25 Primer extension by p66/p51 on RNA335-DNA20. Time course of RNA-dependent DNA synthesis on the random base template by RT is shown on an 8 M urea 10% polyacrylamide gel. Reactions were stopped at 0, 5, 15, 30, 45, 60, 90, 120, 180, and 240 minutes as described under Materials and Methods (lanes 2 - 11). The primer (20-mer) alone (lane 1) was used to indicate the position of the unextended primer. Lanes 12 - 14 are DNA size markers which indicate the position of 25, 31, 33, 34, 35, 44, and 61-mer and were used to determine the position of pause sites.



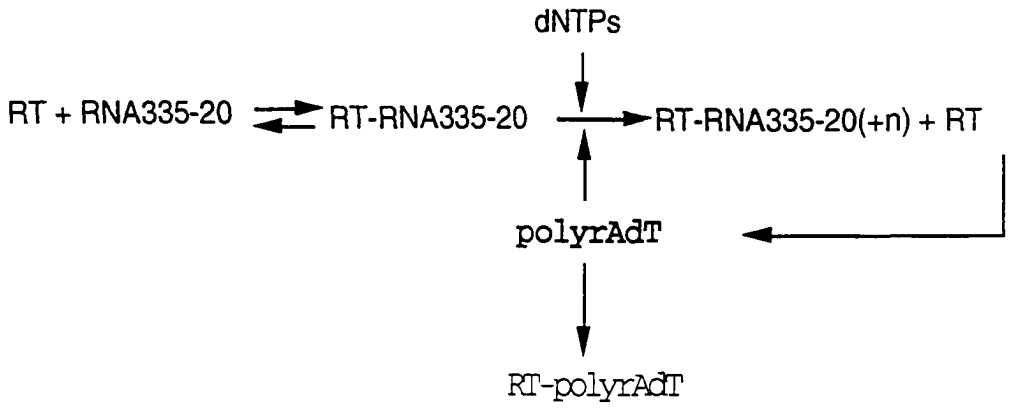
Primer Extension of RT5/6 on RNA335-20

the primer to 62, 65, and 74 nucleotides long and more. Little earlier pauses were detected for p66/p51 like those seen in p64/p51. These results indicate that heterodimer is more processive than homodimer on RNA335-DNA20 template-primer in both wildtype and mutant enzymes. Furthermore, the C-terminal deletion decreases the processivity of the enzyme.

Interaction of HIV-1 RT with RNA335-DNA20

Since a correlation between processivity and the affinity of the enzyme for poly(rA)(dT) was observed, it is important to determine if the processivity is also correlated with the interaction between different RT forms and RNA335-DNA20. The amount of RNA335-DNA20 bound by RT is proportional to the amount of the primer that is extended when RT molecules can be trapped and prevented from binding to the substrate provided that the RT-substrate complex dissociates after a processive cycle of DNA synthesis. A diagram is shown in Figure 26 to depict the method for investigating the interaction between RT and RNA335-DNA20. RT was preincubated with [^{32} P]-labeled RNA335-DNA20 for 10 minutes on ice. Poly(rA)(dT) and dNTPs were added to the mixture to initiate the reaction. The presence of poly(rA)(dT) trapped the free RT molecules and hence prevented the RT molecules from reassociating with RNA335-DNA20. The ability of poly(rA)(dT) to trap free RT molecules was demonstrated by incubating the RT enzyme, RNA335-DNA20, and poly(rA)(dT) first before dNTPs were added to initiate the reaction. The DNA synthesis on RNA335-DNA20 in the presence of poly(rA)(dT) was poorly seen on the autoradiogram as compared to the primer extension results without the presence of poly(rA)(dT). These observation suggest that poly(rA)(dT) is capable of competing the binding of RNA335-DNA20 to the RT molecules. In the binding assays, the bands corresponding to the primer were excised and counted for the amount of unbound RNA335-DNA20.

Figure 26 Scheme to determine Kd for RNA335-DNA20.



The amount of the bound RNA335-DNA20 was obtained by subtracting the total amount of RNA335-DNA20 used by the amount of the unbound RNA335-DNA20. The Eadie plot method was used as described before to obtain the K_d for RNA335-DNA20. The results showed that the K_ds for RNA335-DNA20 were 8.6±0.5, 7.1±0.5, 6.0±0.5, and 18.0±0.5 nM for p64/p64, p64/p51, p66/p66, and p66/p51 respectively as shown in Figure 27 -28. RNA335-DNA20 appeared to bind to p64/p64/, p66/p66, and p64/p51 even tighter than to p66/p51 according to the results, and the processivity on RNA335-DNA20 did not seem to have a significant correlation with the K_ds obtained in these binding assays.

Effect of primer length on affinity for substrate

The affinity for RNA335-DNA12 was examined between wildtype and mutant heterodimers. The RNA-DNA duplex must fit into a primer-template binding groove in RT and contact both DNA polymerase and RNase H active sites. A short primer might cause insufficient interaction with the binding groove and hinder enzyme function. By using the same method described for determining the K_d for RNA335-DNA20, the K_d for RNA335-DNA12 is 27.7 and 2.1 nM for p66/p51 and p64/p51 respectively as shown in Figure 29. The random base RNA template hybridized with a 12-mer was shown to have a weaker interaction with p66/p51 than that hybridized with a 20-mer. The mutant heterodimer has a K_d for RNA335-DNA12 smaller than that for RNA335-DNA20. The random base RNA template-primer appeared to bind tightly to the non-wildtype enzyme as already observed.

Figure 27 Eadie plot of binding assays to determine the affinity of p66/p66 (■-■) and p66/p51 (●-●) for RNA335-DNA20. The standard deviation of each point is less than 5%.

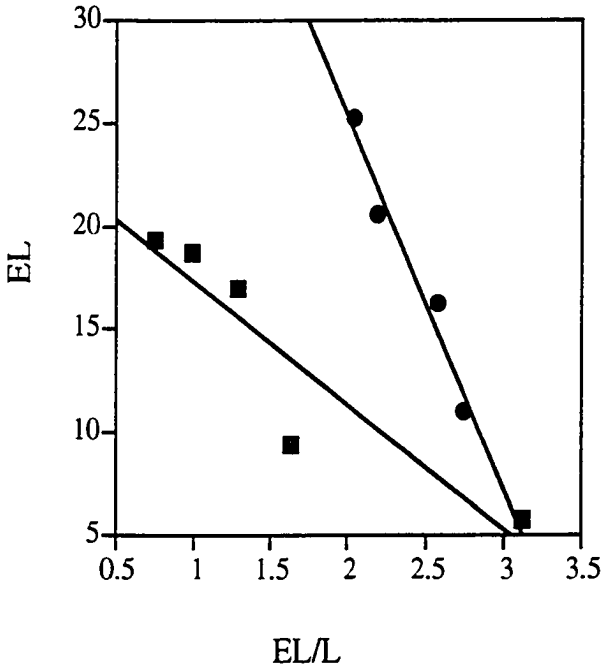


Figure 28 Eadie plot of binding assays to determine the affinity of p64/p664 (■-■) and p64/p51 (●-●) for RNA335-DNA20. The standard deviation of each point is less than 5%.

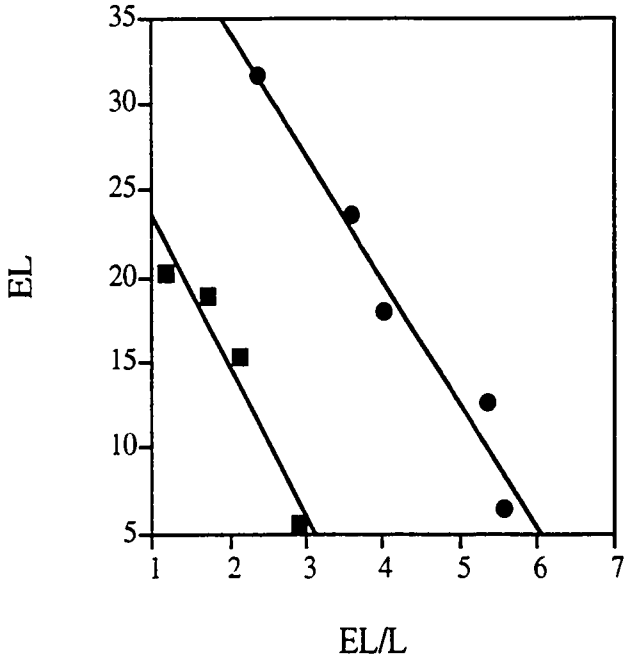
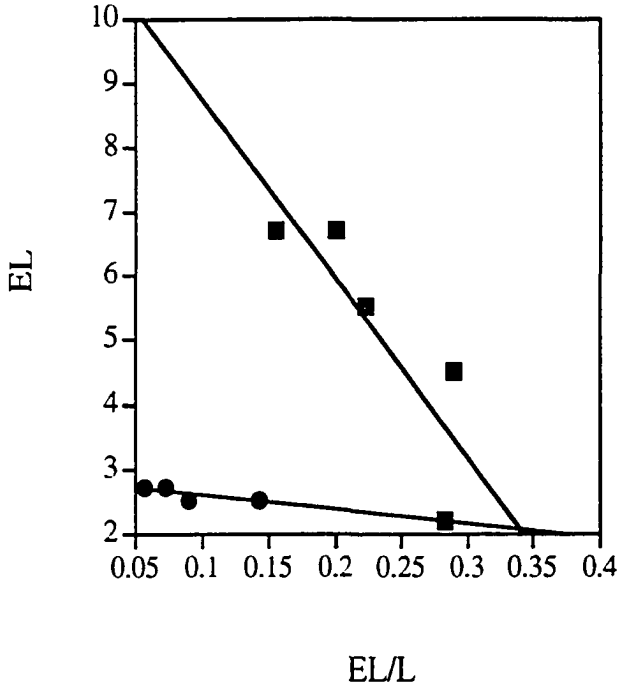


Figure 29 Eadie plot of binding assays to determine the affinity of p66/p51 (■-■) and p64/p51 (●-●) for RNA335-DNA12. The standard deviation of each point is less than 5%.



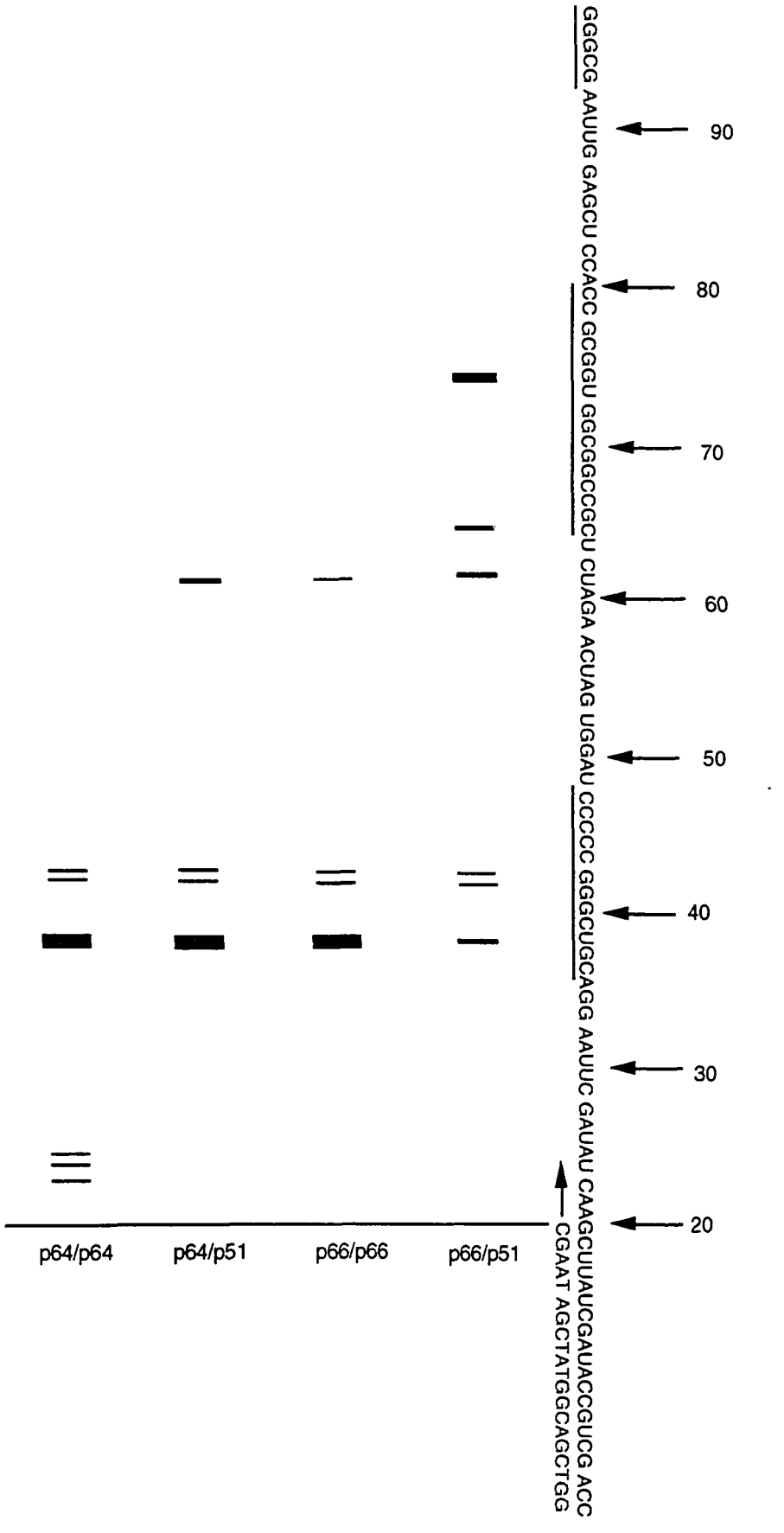
Relationship between the template sequence and pause sites of primer extension

It has been observed that DNA template-directed polymerization by HIV-1 RT is error-prone for single nucleotide substitution, addition and deletion errors in homopolymeric sequences (108). Also, it has been noted that strong termination of processive synthesis occurs at these hot spots (108). It is important to investigate the RNA template sequences, presumably the potential mutation hot spots, where the pauses of processive DNA synthesis tend to occur. As shown in Figures 22-25, DNA synthesis by RT was paused when the primer was extended to specific length. The sequences of the template where the primer extension paused were compared as shown in Figure 30. The template sequences that were related to the pause sites were found to be those rich in G and C bases. The progress of the DNA synthesis by HIV-1 RT was hence dependent on the template sequence. This observation was further confirmed by using two other primers which have the same 5' termini of the 20 nucleotide primer used in the primer extension reaction as showed below :

20-MER	5' GGTCG ACGGT ATCGA TAAGC
12-MER	5' GGTCG ACGGT AT
25-MER	5' GGTCG ACGGT ATCGA TAAGC TTGAT

These three primers were extended to a same pause site (38 nucleotide long) regardless of the length of the primers. Also, the ability to extend over the G C rich template was found various in different form of HIV-1 RT. The wild type heterodimer, p66/p51 was the enzyme most able to extend the primer over the G C rich regions. These results suggest that the processivity of HIV-1 RT on random base template depends on the template sequence. Certain template sequences are able to command RT to pause the DNA synthesis. Also, RT

Figure 30 The sequence alignment of pause sites on random base template by RT.



structure plays an important role in rebinding the primer-template and further extending the terminated primers.

IV. Conclusions and Discussion

In this study, the significance of the p51 subunit and the C-terminal RNase H domain of p66 in RNA-dependent DNA synthesis by HIV-1 RT was investigated. A C-terminal deletion mutant, p64, was used in this investigation. This mutant was shown to have an Ordered Bisubstrate mechanism in its DNA synthesis as in the wildtype enzyme and lacks RNase H activity as described before (97). The deletion includes residues His539 to Leu560 which constitute a loop that connects β sheet 5 to α helix E and the entire α helix itself as shown in Figure 31. Mutations at His539 have been shown to reduce both substrate binding and catalytic rate (105). Also, residues 534 to 545 in HIV-1 RT are the most highly conserved stretch of 12 amino acids in five lentivirus RNase H sequences as shown in Figure 32 (105). According to the crystal structure, p66 appears to be the catalytic subunit by providing a cleft for substrate binding and catalysis sites for DNA polymerase and RNase H activity. However, p51 assumes a different structure and is the non-catalytic subunit (50). The RNase H domain of p66 is responsible for hydrolyzing the RNA template on RNA-DNA duplex during DNA synthesis. Physical contact between the RNase H domain and the primer-template was suggested by computer modeling of HIV-1 RT and RNA-DNA interaction (50). It is hence important to understand the effects of the deletion of these conserved residues on the processivity and primer-template binding.

Figure 31 C-terminal deletions in the RNase H domain of HIV-1 RT. The amino acid sequence and the secondary structure assignments for the RNase H domain are shown according to the crystal structure (98). The amino acid deleted in the mutant enzyme (p64 or RT4) is underlined.

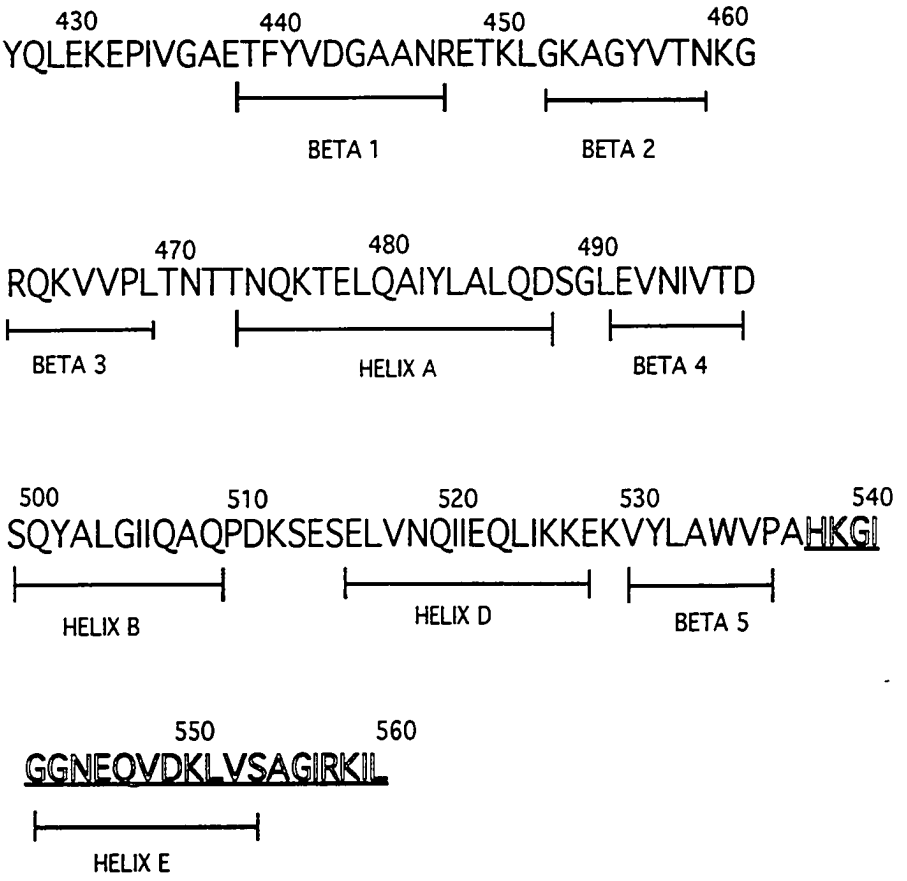


Figure 32 The deletion in the mutant RT contains part of the conserved residues present in RNase H domains of lentiviruses. (HIV-2, human immunodeficiency virus type 2; SIV, simian immunodeficiency virus from macaque monkey; EIVA, equine infectious anemia virus; VISNA, visna lentivirus) These conserved residues are thought to cluster near the catalytic site (98).

	534	545
HIV-1	AWVPAHKGIGGN	
HIV-2	AWVPAHKGIGGN	
SIV	AWVPAHKGIGGN	
EIAV	AWVPGHKGITGY	
VISNA	AWVPGHKGIPPN	

guided by the direct repeat sequence (R) on both ends of the RNA, and minus strand DNA synthesis proceeds to the 5' end of the RNA (PBS). The primer for plus strand DNA synthesis is derived from the RNA genome at a highly conserved region called the polypurine tract (PPT), which is located near the U3 region at the 5' end of the minus strand DNA. The synthesis of the plus strand DNA is initiated from PPT and continued to the PBS of tRNA. This DNA plus strand is then translocated to the 3' end of the minus strand, guided by the PBS sequences at 3' end of both plus and minus strands. The tRNA and the primer for plus strand synthesis are subsequently removed by a precise RNase H cleavage that defines the right and left ends of the LTRs, respectively (55-57, 59-62). The sequences of the LTRs are important for the integration of the viral DNA into the host cell chromosome. The LTRs also contain promoter, enhancer, and termination signals for gene expression. The synthesis of viral DNA is complete after both strands are extended to the 5' end of the complementary strand. The RNase H activity associated with HIV-1 RT is involved in selective degradation of RNA in the DNA hybrid and in generation of an RNA primer for synthesis of plus strand DNA (25, 43, 63-65). It is also responsible for specific removal of the host tRNA primer as well as the plus strand primer by specific cleavage at RNA-DNA junction.

The residues Pro1 through Tyr266 of HIV-1 RT are thought to be required for polymerase domain (47). Asp 185 and Asp186 are required in the polymerase active site (50). A bireactant - biproduct mechanism has been proposed for DNA synthesis by HIV-1 RT (66, 67). In this mechanism, binding of the reactants is ordered with the primer-template binding first followed by the binding of the 2'-deoxynucleoside 5'-triphosphate (dNTP) that is complementary to the next appropriate template residue as shown in Figure 5. The incorporation of the dNMP into the nascent chain then occurs via an S_N2 displacement of

The dissociation constant (K_d) for poly(rA)(dT) obtained by nitrocellulose binding assays suggested that the processivity of HIV-1 RT on poly(rA)(dT) correlates with the affinity for the substrate as summarized in Table 2. These results support the hypothesis that p51 acts to enhance the processivity of DNA synthesis by improving the affinity for primer-template. It is conceivable that having an RNase H domain on both subunits as in the p66/p66 and p64/p64 homodimers can cause interference in primer-template binding.

It has been proposed that the processivity of a polymerase comes from a tight interaction between enzymes and certain repeating motifs in the substrate; distributive enzymes bind to the motifs that do not repeat (103). Also, it has been reported that processivity numbers of HIV-1 RT can vary widely with the template used, increasing from a few to > 300 nucleotides (104). It is hence important to examine the enzyme function on templates other than poly(rA), particularly those with random bases instead of repeating the same base. A 335 nucleotide random base RNA template was then synthesized for the further studies. This template was hybridized with a 20 nucleotide primer (Figures 10 and 11). Processivity on this RNA335-DNA20 was seen to be affected by both the enzyme structure and pause sites as observed on the autoradiograms (Figures 22-25). In Figure 22, strong pause sites at 23, 24, and 38 nucleotides by p64/p64 mutant homodimer were observed. This C-terminal deletion mutant homodimer, which has a defect in RNase H function, a physical deletion in the polypeptide , and extra 98 residues in the non-catalytic subunit, was distributive for the first 3-5 nucleotides incorporation. In p64/p51 (Figure 23), primers were extended to 38 nucleotides before any major pause sites were observed. The major pause site at 38 nucleotides was also observed in p66/p66 as in p64/p51. However, p64/p51 was seen to be able to further extend the primer to 62 nucleotides. This observation indicates that p64/p51 is more active than p66/p66 in terms of DNA

synthesis on RNA335-DNA20. The extra RNase H domain was seen to cause hindrance when comparing the primer-extension pattern by p64/p51 heterodimer to those by p66/p66 and p64/p64 homodimer. Also, p51 was seen to contribute to processivity when the primer extension patterns by p66/p66 and p66/p51 are compared. Although, the pause site at 38 nucleotides was also observed in p66/p51, the longer DNA products (74 nucleotides and more) predominated. The order of processivity on RNA335-DNA20 was: p66/p51> p64/p51> p66/p66> p64/p64. The enzyme structure was shown to play an important role in processivity on RNA335-DNA20 as on poly(rA)(dT). This dependence of processivity on enzyme structure was further studied by examining the interaction between RNA335-DNA20 and the different forms of enzyme.

As described above, the increases in enzyme activity and processivity on poly(rA)(dT) are correlated with the interaction between RT enzyme and poly(rA)(dT). Alteration in enzyme structure cause changes in enzyme affinity for poly(rA)(dT) and hence influence the enzyme activity and processivity. It is conceivable that the less processive RT molecules associate with lower affinity for RNA335-DNA20. However, when K_ds were examined (Figures 27 and 28), an opposite result was obtained. The p66/p51 heterodimer maintains a Kd for RNA335-DNA20 similar to that for poly(rA)(dT). But, p66/p66, p64/p64, and p64/p51 appeared to interact with RNA334-DNA20 stronger than p66/p51 and had higher affinity for this random base template-primer than that for poly(rA)(dT). Since the binding assays to determine the K_d for RNA335-DNA20 depends on the ability of poly(rA)(dT) to trap the free RT, and poly(rA)(dT) was already shown to have different affinity for different forms of RT, the K_ds for RNA335-DNA20 observed might be affected by the failure of poly(rA)(dT) to trap all the free RT molecules. For example, the p64/p64 has a K_d of 392 nM for poly(rA)(dT), and the use of poly(rA)(dT) to trap the free p64/p64 could have

failed to prevent all the p64/p64 molecules from reassociating with RNA335-DNA20. This problem was prevented by using excess of poly(rA)(dT) to favor the formation of RT-poly(rA)(dT) complex. In a control reaction where RT, poly(rA)(dT), and RNA335-DNA20 were incubated at the same time before the reaction was initiated, the amount of poly(rA)(dT) used was shown to be able to trap RT. The [^{32}P]-labeled primer in this control reaction was poorly extended as judged by the autoraiogram. The tight interaction between random RNA template-primer and the non-wildtype enzyme was further observed when RNA335-DNA12 was used to study the effect of primer length on enzyme affinity for substrate (Figure 29). As summarized in Table 3, the non-wildtype RT molecules have smaller K_ds for random base RNA template-primer than the p66/p51 wildtype enzyme.

Table 3 Summary of the measured dissociation constants (nM) of HIV-1 RT for random base RNA template-primer .

template-primer	p66/p51	p64/p51	p64/p64	p66/p66
RNA335-DNA20	18.0±0.5	7.1±0.5	8.6±0.5	6.0±0.5
RNA335-DNA12	27.7±0.5	2.1±0.5		

The p64/p51 molecules appeared to have stronger interaction with RNA335-DNA12 than RNA335-DNA20. However, the wildtype enzyme, p66/p51, demonstrated a weaker affinity for RNA335-DNA12 than for RNA335-DNA20.

This result was similar to that observed by Beard (74). In his studies, short oligo (dT) primers (less than 14 nucleotides), annealed to poly(rA), bound RT relatively weakly (K_d = 20 - 30 nM), whereas oligo (dT) primers of 16 and 20 nucleotides in length bound tightly to RT (K_d ~ 0.1 nM). The results of the binding assays suggest the processivity on RNA335-DNA20 does not depend on the dissociation constant of RT for this template-primer. The tight interaction observed could be because the substrate binding in those non-wildtype enzymes is more sensitive to the nature of the primer-template. Strong premature termination was seen in the primer extension on RNA335-DNA20 by the non-wildtype enzyme. This observation suggests that these RT molecules, after binding and short extending the primer, either dissociate from and lose affinity for the primer-template or remain bound to the primer-template without any further extension. A "stall" mechanism has been proposed to explain the premature termination in DNA synthesis by HIV-1 RT (78). This mechanism suggested that the enzyme remains bound strongly to the primer-template and allows no further DNA synthesis.

The pauses of DNA synthesis tended to occur at template positions containing poly G,C sequences as shown in Figure 30. This dependency of pause sites on template sequence was confirmed by using a 12-mer and a 25-mer as primer which have the same 5' termini of the 20-mer. The major pause site, 38-mer, was seen to be independent of the position of the 3' end of the primer but dependent on the template sequence. The template sequence flanking the pause sites are shown in Figure 33. The G-C sequence dependence of the pause sites is an important area to study to elucidate why HIV-1 RT can catalyze synthesis over all the G-C rich stretches in its RNA genome to complete the transcription. Mutations might be required for RT to transcribe over these G-C rich pause sites. Furthermore, an auxiliary viral or cellular factor might be

Figure 33 The template sequences around the pause sites for DNA synthesis by HIV-1 RT on RNA335-DNA20. The length of the paused product and seven bases upstream and down stream the pause site are shown. The poly G or C stretches are underlined.

PAUSE SITES	SS template 7 6 5 4 3 2 1	dS Stem 0 1 2 3 4 5 6 7
20	A U A U C A A	**G** C U U A U C G
23	U G C A U A U	**C** A A G C U U A
24	U U G C A U A	**U** C A A G C U U
25	A U U G C A U	**A** U C A A G C U
38	C C G G G C U	**G** C A G G A A U
43	A U C C C C C	**G** G G C U G C A
44	G A U C C C C	**C** G G G C U G C
62	G C C G C U C	**U** A G A A C U A
65	G C G G C C G	**C** U C U A G A A
74	A C C G C G G	**U** G G C G G C C

5'

involved in the process. The G-C rich stretches are likely forming secondary structure through intra- or intermolecular interaction between G and C bases and hence blocking the DNA synthesis. It is also possible that HIV-1 RT contains substrate binding subsites for the RNA template, similar to those known for RNase A and that these subsites happen to have strong affinity for G and C bases. If this hypothetical subsite scenario is correct, then polymerization proceeds across the RNA template in a processive fashion through AT rich areas. However, when HIV-1 RT encounters several G and C bases in a row, most of the subsites interact with G or C and hence lock the primer-template. The DNA synthesis is then terminated by the "stall" mechanism described above. On the other hand, it is also possible that the subsites have very loose affinity for G and C and the termination of DNA synthesis is a consequence of the loose interaction between enzyme and G-C rich areas of the primer-template. HIV-1 RT falls off the primer-template and fails to extend the primer over the G-C rich template.

The processivity data of p64/p51 and p66/p51 showed that the C-terminal deletion can reduce the processivity of DNA synthesis on both poly(rA)(dT) and RNA335-DNA20. Since the mutant enzyme still maintains its Ordered Bisubstrate mechanism in DNA synthesis, this 22 amino acid deletion in the RNase H domain does not affect the assembly of the functional polymerase domain, but it is important in processivity. Similar results were observed in a Gln475Glu mutant of HIV-1 RT which has retarded RNase H function (107). In the primer extension study, the primers were not extended by p64/p51 as well as they were by p66/p51. A catalytically functional RNase H domain is hence required to facilitate processive DNA synthesis and reinitiation of the DNA synthesis at specific pause sites. The role of the RNase H activity in facilitating DNA synthesis was also observed by Dudding *et al.* (90). However, as seen in

the processivity and primer extension data by p66/p66, although a functional RNase H domain was present in the homodimer, the processivity was still reduced and the primer was not well extended. It is likely the extra RNase H domain in the non-catalytic subunit causes hindrance to DNA synthesis. It is also possible that the folding of the non-catalytic p66 is less capable of maintaining the favorable enzyme conformation than p51 in the wildtype.

It is hoped that the knowledge of structure-function relationships of HIV-1 RT will facilitate the design of new inhibitors of this enzyme and suggest strategies that might help to circumvent the ability of mutant RT molecules to escape the potency of inhibitors. In this investigation, the p51 subunit is seen to contribute significantly to enzyme activity and processivity of HIV-1 RT. The extra RNase H domain in the small subunit can decrease the enzyme function, as seen in p66/p66 and p64/p64. The p51 subunit, although harbors no catalytic sites, plays an important role in maintaining the favorable conformation of the enzyme. Hence, molecules that can disrupt the interaction between subunits are potential HIV-1 RT inhibitors. The C-terminal deletion in p64/p51 causes decreases in enzyme activity and processivity. Since the HIV-1 RT DNA polymerase activity is correlated with its RNase H domain, inhibitors that can affect the RNase H function or alter its conformation can be useful in blocking the DNA synthesis. G-C rich sequences of RNA template is shown to be able to command the RT to pause the DNA polymerization. Introduction of molecules that can prevent the RT from reinitiating the DNA synthesis on these pause sites is hence one of the antiviral strategies.

V. References

1. Barre-Sinousi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chameret, S., Gruest, J., Dauget, C., Axler-Blin, C.,Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., and Montagnier, L. (1983) Science **220**, 868-871.

2. Levy, J. A., Hoffman, A. D., Kramer, S. M., Landin, J. A., Schimabukaro, J. M., and Oshiro, L.S. (1984) Science **225**, 840-842.

3. Popovic, M., Sarngadharan, M. G., Read,E., and Gallo, R. C. (1984) Science **224**, 497-500.

4. Fauci, A. S. (1988) Science **239**, 617-622.

5. Gallo, R. C. and Montagnier, L. (1988) in The Science of AIDS, 1-11, W. H. freeman and Co., New York.

6. Mitsuya, H., and Broder, S. (1987) Nature **325**, 773-778.

7. Take, Y., Inouye, Y., and Nakamura, S. (1989) J. Antibiot. **42**,107-115.

8. Goldman, M. E., Salituro, G. S., Bowen, J. A., Williamson, J. M., Zink, D. L., Schleif, W. A., and Emini, E. A. (1990) Mol. Pharmacol. **38**, 20-25.

9. Nakane, H., and Ono, K. (1990) Biochemistry **29**, 2841-2845.

10. Ono, K., Nakane, H., Fukushima, M., Chermann, J. C., and Barre-Sinousi, F. (1990) Eur. J. Biochem. **190**, 469-476.

11. Schatz, O., Cromme, F.,Grueninger-Leitch, F., and Le Grice, S. F. J. (1989) FEBS Lett. **257**, 311-314.

12. Tan, G. T., Kinghorn, A. D., Hughes, S. H., and Pezzuto, J. M. (1991) J. Biol. Chem. **266**, 23529-23536.

13. Merluzzi, V. J., Hargrave, K. D., Labadia, M., Grozinger, K., Skoog, M., Wu, J. C., Shih, C., Eckner, K., Hattox, S., Adams, J., Rosehthal, A. S., Faanes, R., Eckner, R. J., Koup, R. A., and Sullivan, J. L. (1990) Science **250**, 1411-1413.

14. Cheng, Y., Gao, W., Chen, C., Vazquez-Padua, M., and Starnes, M. C. (1991) Annals N.Y. Acad. Sci. **616**, 217-222.

15. Cheng, Y., Dutschman, G. E., Bastow, K. F., Sarngadharan, M. G., and Ting, R. Y. C. (1987) J. Biol. Chem. **262**, 2187-2189.

16. Parker, W. B., White, E. L., Shaddix, S. C., Ross, L. J., Buckheit, Jr.,R. W., Germany, J. M., Secrist III, J. A., Vince, R., and Shannon, W. M. (1991) J. Biol. Chem. **266**, 1754-1762.

17. Baba, M., Clercq, E. D., Tanaka, H., Ubasawa, M.,Takashima, H., Sekiya, K., Nitta, I., Umezu, K., Nakashima, H., Mori, S., Shigeta, S., Walker, R. T., and Miyasaka, T. (1991) Proc. Natl. Acad. Sci. USA **88**, 2356-2360.

18. Mitsuya, H., Yarchoan, R., Kageyama, S., and Broder, S. (1991) FASEB J. **5**, 2369-2381.

19. Wu, J. C., Warren, T. C., Adams, J., Proudfoot, J., Skiles, J., Raghavan, P., Perry, C., Potocki, I., Farina, P. R., and Grob, P. M. (1991) Biochemistry **30**, 2022-2026.

20. Larder, B.A., and Kemp, S.P. (1990) Science **246**, 1155-1158.

21. Land, S., Treloar, G., McPhee, D., Birch, C., Doherty, R., Cooper, D. and Gust, I. (1990) J. Infect. Dis. **161**, 326-329.

22. Larder, B. A., Darby, G. and Richman,D. D. (1989) Science **243**, 1731-1734.

23. Haseltine, W. A. and Wong-Staal, F. (1988) in The Science of AIDS ,12-25, W. H. Freeman and Co., New York.

24. 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., Ivanoff, L., Petteway Jr, S. R., Pearson, M. C., Lautenberger, J. A., Papas, T. S., Ghrayeb, J.,Chang, N. T., Gallo, R. C., and Wong-Staal, F. (1985) Nature **313**, 277-284.

25. Starnes, M. C., Gao,W., Ting, R. Y. C., and Cheng, Y. (1988) J. Biol. Chem. **263**, 5132-5134.

26. Di Marzo Veronese, F. M., Copeland, T. D., Devoice, A. L., Rahman, R., Oroszlan, S., Gallo, R. C., and Sarngadharan, M. G. (1986) Science **231**, 1289-1291.

27. Chandra, A., Gerber, T., and Chandra, P. (1986) FEBS Lett. **197**, 84-88.

28. Muller, B., Restle, T., Weiss, S., Gautel, M., Sczakiel, G., and Goody, R.

(1989) J. Biol. Chem. **264**, 13975-13978.

29. Lightfoote, M. M., Coligan, J. E., Folks, T. M., Fauci, A. S., Martin, M. A., and Venkatesan, S. (1986) J. Virol. **60**, 771-775.

30. Farmerie, W. G., Loeb, D. D., Gasavant, N. C., Hutchinson, C. A., Edgell, M. H. and Swanstrom, R. (1987) Science **236**, 305-308.

31. Mizrahi, V., Lazarus, G. M., Miles, L. M., Meyers, C. A. and Debouck, C. (1989) Arch. Biochem. Biophys. **273**, 347-358.

32. Le Grice, S. F. J., Mills, J., Ette, R., and Mous, J. (1989) J. Biol. Chem. **264**, 14902-14908.

33. Graves, M. C., Meidel, M. C., Pan, Y. C. E., Manneberg, M., Lahm, H. W. and Gruninger-Leitch, F. (1990) Biochem. Biophys. Res.Commun. **168**, 30-36.

34. Becerra, S. P., Clore, G. M., Gronenborn, A. M., Karlstrom, A. R., Stahl, S. J., Wilson, S. H. and Wingfield, P. T. (1990) FEBS. Lett. **270**, 76-80.

35. Bathurst, I. C., Moen, L. K., Lujan, M. A., Gibson, H. C., Feucht, P. H., Pichuantes, S., Craik, C. S., Santi, D. V. and Barr, P. J. (1990) Biochem. Biophys. Res. Commun. **171**, 589-595.

36. Johnson, M. S., McClure, M. A., Feng, D. F., Gray, J. and Doolittle, R. (1986) Proc. Natl. Acad. Sci. USA **83**, 7648-7652.

37. Larder, B., Purifoy, D.J.,Powell, K. L., and Darby, G. K. (1987) EMBO J. **6**, 3133-3137.

38. Hizi, A., McGill, C., and Hughes, S. H. (1988) Proc. Natl. Acad. Sci. USA **85**, 1218-1222.

39. Hizi, A., Barber, A. and Hughes, S. H. (1988) Virology **170**, 326-329.

40. Prasad, V. R. and Goff, S. P. (1989) Proc. Natl. Acad. Sci. USA **86**, 3104-3108.

41. Lowe, D. M., Aitken, A., Bradley, C., Darby, G. K., Larden, B. A., Powell, K. L., Purifoy, D. J. M., Tisdale, M., and Stammers, D. K. (1988) Biochemistry **27**, 8884-8889.

42. Tisdale, M., Ertl, P., Larder, B. A., Purifoy, D. J. M., Darby, G., and Powell, K. L. (1988) J. Virol. **62**, 3662-3667.

43. Hansen, J., Schulze, T., Mellert, W., and Molling, K. (1988) EMBO J. **7**, 239-243.

44. Tanese, N. and Goff, S. P.(1988) Proc. Natl. Acad. Sci. USA **85**, 1777-1781.

45. Poch, O., Sauvaget, I., Delarue, M., and Tordo, N. (1989) EMBO J. **8**, 3867-3874.

46. Ferris, A. L., Hizi., A., Showalter, S. D.,Pichuantes, S., Babe, L., Craik, C. S., and Hughes, S. H. (1990) Virology **175**, 456-464.

47. Jacobo-Molina, A. and Arnold, E. (1991) Biochemistry **320**, 6351-6361.

48. Painter, G. R., Wright, L.L., Hopkins, S. and Furman, P. A. (1990) J. Cellular Biochem. **14D**, 109.

49. Cheng, N.,Painter, G. R. and Furman, P. A. (1991) Biochem. Biophys. Res. Commun. **174**, 785-789.

50. Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1992) Science **256**, 1783-1790.

51. Temin, H. M., and Mizutani, S. (1970) Nature **226**, 1211-1213.

52. Baltimore, D. (1970) Nature **226**, 1209-1211.

53. Moelling, K., Bolognesi, D., Bauer, H., Busen, W., Plassmann, W., and Hausen, P. (1971) Nature New Biol. **234**, 240-244.

54. Collett, M. S.,Dierks, P., Parson, J. T., and Faras, A. J. (1978) Nature **272**, 181-184.

55. Gilboa, E., Mitra, S. W., Goff, S., and Baltimore, D. (1979) Cell **18**, 93-100.

56. Smith, J. K., Cywinski, A., and Taylor, J. M. (1984) J. Virol. **49**, 200-204.

57. Resnick, R., Omer, C. A., and Faras, A. J. (1984) J. Virol. **51**, 813-821.

58. Wain-Hoboson, S., Sonigo, P., Danos, O., and Alizon, M. (1985) Cell **40**, 9-17.

59. Champoux, J. J., Gilbou, E., and Baltimore, D. (1984) J. Virol. **49**, 686-691.

60. Finston, W. I., and Champoux, J. J. (1984) J. Virol. **51**, 26-33.

61. Smith, J. K., Cywinski, A., and Taylor, J. M. (1984) J. Virol. **51**, 314-319.

62. Huber, H. E. and Richardson,C. C. (1990) J. Biol. Chem. **265**, 10565-10573.

63. Krug, M. S., and Berger, S. L. (1989) Proc. Natl. Acad. Sci. USA **86**, 3539-

3543.

64. Oyama, F., Kikuchi, R., Crouch, R. J., and Uchida, T. (1989) J. Biol. Chem. **264**, 18808-18817.

65. Gerard, G. F. (1981) Biochemistry **20**, 256-265.

66. Majumdar, C., Abbotts, J.Broder, S., and Wilson, S. H. (1988) J. Biol. Chem. **263**, 15657-15665.

67. Huber, H. E., McCoy, J. M., Seehra, J. S., and Richardson, C. C. (1989) J. Biol. Chem. **264**, 4669-4678.

68. Wu, J. C., Warren, T. C., Adams, J. ,Proudfoot, J., Skiles, J., Raghavan, P., Pevry, C., Potocki, I., Farina, P. R., and Grob, P. M. (1991) Biochemistry **30**, 2022-2026.

69. Basu, A., Ahluwalia, K. K., Basu, S., and Modak, M. J. (1992) Biochemistry **31**, 616.

70. Kumar, A., Kim, H., Sobol, R. W., Becerra, S. P., Lee, B., Hatfield, D. C., Suhadolnik, R. J., and Wilson, S. H. (1993) Biochemistry **32**, 7466-7474.

71. Reardon, J. E. (1992) Biochemistry **31**, 4473-4479.

72. Kati, W. M. and Johnson, K. A. (1992) J. Biol. Chem. **267**, 25988-25997.

73. Hsieh, J., Zinnen, S., and Modrich, P. (1993) J. Biol. Chem. **268**, 24607-24613.

74. Beard, W. A. and Wilson, S. H. (1993) Biochemistry **32**,9745-9753.

75. Bryant, F. R., Johnson, K. A., and Benkovic, S. J. (1983) Biochemistry **22**, 3546-3551.

76. Abbots, J., Bebenek, K., kunkel, T. A., and Wilson, S. H. (1988) J. Biol. Chem. **268**, 10312-10323.

77. Abbots, J., SenGupta, D.V., Zon, G., and Wilson, S. H. (1988) J. Biol. Chem. **263**, 15094-15103.

78. Klarmann, G. T., Schauber, C. A., and Preston, B. D. (1993) J. Biol. Chem. **268**, 9793-9802.

79. Kim, S., Byrn, R., Groopman, J., and Baltimore, D. (1989) J. Virol. **63**, 3708-3713.

80. Sato, H., Orenstein, J. ,Dimitrov, D., and Martin, M. (1992) Virology **180**, 712-724.

81. Williams, K. J., Loeb, L. A., and Fry, M. (1990) J. Biol. Chem. **265**, 18682-18689.

82. Kornberg, A., and Baker, T. A. (1992) DNA Replication, 2nd Ed., W. H. Freeman and Co., New York.

83. Hasen, J., Schulze, T., Mellert, W., and Moelling, K. (1988) EMBO J. **7**, 239-243.

84. Hostomsky, Z., Hostomsky, Z., Hudson, G. Q., Moomaw, E. W., and Nodes, B. R. (1991) Proc. Natl. Acad. Sci. USA **88**, 1148-1152.

85. Hu, S. C., McKelvin, D., Zweig, M., Court, D. L., and Gerwin, B. I. (1988) J. Virol. **62**, 4376-4380.

86. Schatz, O., Mous, J., and Le Grice, S. F. J. (1990) EMBO J. **9**, 1171-1176.

87. DeStefano, J. J., Biser, R. G., Mallaber, L. M., Myers, T. W., Bambara, R. A. and Fay, P. J. (1991) J. Biol. Chem. **266**, 7423-7431.

88. DeStafano, J. J., Buiser, R. G., Mallaber, L. M., Bambara, R. A., and Fay, P. J. (1991) J. Biol. Chem. **226**, 24295-24301.

89. Volkmann, S., Wohrl, B., Tisdale, M., and Moelling, K. (1993) J. Biol. Chem. **268**, 2674-2683.

90. Dudding, L. R., Nkakinde, N. C., and Mizrahi, V. (1991) Biochemistry **30**, 10498-10506.

91. Dudding, L. R. and Mizrahi, V. (1993) Biochemistry **32**, 6116-6120.

92. Bavand, M. R., Wagner, R., and Richmond, T. J. (1993) Biochemistry **32**, 10543-10552.

93. Restle, T., Mueller, B.,and Goody, R. S. (1990) J. Biol. Chem. **265**, 8986-8988.

94. Howard, K. J., Frank, K. B., Sim, I. S., and Legrice, S. F. J. (1991) J. Biol. Chem. **266**, 23002-23009.

95. Barr, P. J., Power, M. D., Lee-Ng, C. T., Gibson, H. L., and Lucin, P. A. (1987) Bio/Technology **5**, 486-489.

96. Moen, L. K., Bathurst, I. C., and Barr, P. J. (1992) AIDS Research and Human Retroviruses **8**, 597-604.

97. Heard, T. S. (1992) Kinetic Characterization of a Recombinant C-terminal Mutant of Reverse Transcriptase from the Human Immunodefficiency virus, M.S. Thesis, Old Dominion Univ..

98. Davies, J. F., Hostomska, Z., Hostomsky, Z., Jordan, S. R.,and Matthew, D. A. (1991) Science **252**, 88-93.

99. Furfine, E. S., and Reardon, J. E. (1991) J. Biol. Chem. **266**, 406-412.

100. Fersht, A. (1985) Enzyme Structure and Mechanism, 191-192, W. H. Freeman and Co., New York.

101. Lewis, M. S. (1991) Biochemistry **30**, 11707-11719.

102. Huang, S. C., Smith, J. R., and Moen, L. K. (1992) Biochem. Biophys. Res. Comm. **184**, 986-992.

103. del Cardayre, S. B. and Raines, R. T. (1994) Biochemistry **33**, 6031-6037.

104. Reardon, J. E. (1993) J. Biol. Chem. **268**, 8743-8751.

105. Davies, J. F., Hostomska, Z., Hostomsky, Z., Jordan, S. R., Matthew, D. A. (1991) Science **252**, 88-95.

106. Bavand, M. R., Wagner, R., and Richmond, T. J. (1993) Biochemistry **32**, 10543-10552.

107. Volkmann, S., Wohrl, B. M., Tisdale, M., and Moelling, K. (1993) J. Biol. Chem. **268**, 2674-2683.

108. Bebenek, K., Abbotts, J., Roberts, J. D., Wilson, S. H., and Kunkel, T. A. (1989) J. Biol. Chem. **264**, 16948-16956.

APPENDIX 1

A. pBluescript (+) sequence for generating the 446 bp DNA fragment

530 540 550 560 570

GCTG GCGAA GGGGG GATGT GCTGC AAGGC GATTA AGTTG GGTAA CGCCA

PCR primer 1

580 590 600 610 620

GGGTT TTCCC AGTCA CGACG TTGTA AAACG ACGCC CAGTG AATTG TAATA

T7

630 640 650 660 670

CGACT CACTA TAGGG CGAAT TGGAG CTCCA CCGCG GTGGC GGCCG CTCTA

Promoter **+1**

680 690 700 710 720

GAACT AGTGG ATCCC CCGGG CTGCA GGAAT TCGAT ATCAA GCTTA TCGAT

730 740 750 760 770

ACCGT CGACC TCGAG GGGGG GCCCG GTACC CAGCT TTTGT TCCCT TTAGT

780 790 800 810 820

GAGGG TTAAT TCCGA GCTTG GCGTA ATCAT GGTCA TAGCT GTTTC CTGTG

830 840 850 860 870

TGAAA TTGTT ATCCG CTCAC AATTC CACAC AACAT ACGAG CCGGA AGCAT

880 890 900 910 920

AAAGT GTAAA GCCTG GGGTG CCTAA TGAGT GAGGT AACTC ACATT AATTG

930 940 950 960 970

CGTTG CGCTC ACTGC CCGCT TTCCA GTCGG GAAAC CTGTC GTGCC AG

GCC CTTTG GACAG CACGG TC

PCR primer 2

APPENDIX 2

B. Sequence of the random base RNA template

```
          10          20          30          40          50
           *           *           *           *           *
GGGCG AAUUG GAGCU CCACC GCGGU GGCGG CCGCU CUAGA ACUAG UGGAU

          60          70          80          90         100
           *           *           *           *           *
CCCCC GGGCU GCAGG AAUUC GAUAU CAAGC UUAUC GAUAC CGUCG ACCUC
                                  CG AATAG CTATG GCAGC TGG 5'
                                       20-MER

         110         120         130         140         150
           *           *           *           *           *
GAGGG GGGGC CCGGU ACCCA GCUUU UGUUC CCUUU AGUGA GGGUU AAUUC

         160         170         180         190         200
           *           *           *           *           *
CGAGC UUGGC GUAAU CAUGG UCAUA GCUGU UUCCU GUGUG AAAUU GUUAU

         210         220         230         240         250
           *           *           *           *           *
CCGCU CACAA  UUCCA CACAA CAUAC GAGCC GGAAG CAUAA AGUGU AAAGC

         260         270         280         290         300
           *           *           *           *           *
CUGGG GUGCC UAAUG AGUGA GGUAA CUCAC AUUAA UUGCG UUGCG CUCAC

         310         320         330
           *           *           *
UGCCC GCUUU CCAGU CGGCA AACCU GUCGT GCCAG
```