

https://dspace.org/

The role of the conserved ASP443 and ASP498 residues in the polymerase and RNase H activities of HIV-1 reverse transcriptase

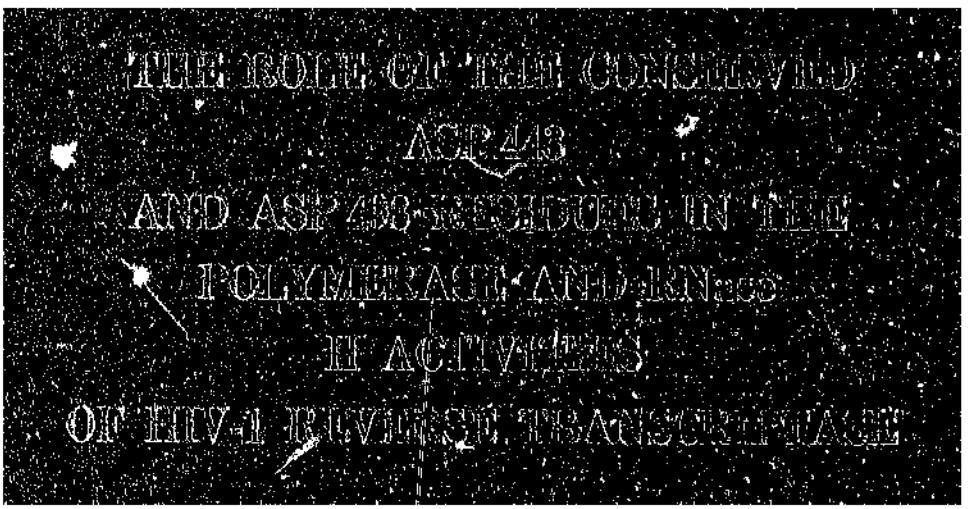
Brooksbank, Richard L

1993

http://hdl.handle.net/10539/21434

Brooksbank, Richard L. (1993) The role of the conserved ASP443 and ASP498 residues in the polymerase and RNase H activities of HIV-1 reverse transcriptase, University of the Witwatersrand, Johannesburg, <http://wiredspace.wits.ac.za/handle/10539/21434>

Downloaded from a DSpace repository. https://dspace.org/


THE ROLE OF THE CONSERVED
ACID
AND ASP 498 RESIDUES IN THE
POLYMERASE AND RNase
H ACTIVITIES
OF HIV-1 REVERSE TRANSCRIPTASE

THE ROLE OF THE CONSERVED ASP443 AND ASP498 RESIDUES IN THE POLYMERASE AND RNase H ACTIVITIES OF HIV-1 REVERSE TRANSCRIPTASE.

Richard L. Brooksbank

Degree awarded with distinction on 8 Dec 1993.

Submitted in fulfilment of the requirements for the degree of Master of Science in the faculty of Science, University of the Witwatersrand, Johannesburg.

Johannesburg 1993.

ABSTRACT.

The roles of the highly conserved aspartic acid residues found at positions 443 and 498 within the RNase H domain of Human Immunodeficiency Virus type-1 reverse transcriptase (RT) were investigated by the defined substitution of these residues using site-directed mutagenesis. Two mutant RT enzymes were created using this technique; an Asp443 to Ala443 (D443A) mutant and a double mutant, Asp443 to Ala443 and Asp498 to Asn498 (D443A/D498N). Authentic HIV-1 RT was obtained by co-expression of the mutant enzymes with HIV-1 protease followed by purification. Subsequent characterisation of the enzymes revealed that the D443A RT mutant lacked a RNase H activity but retained normal RNA-dependent DNA polymerase activity relative to that of the wild-type enzyme when using a homopolymer RNA.DNA substrate. An assessment of the thermal stability of the D443A RT enzyme showed that the enzyme was as stable as wild-type HIV-1 RT. The D443A/D498N mutant was not recoverable from the co-expression system, an indication that the substitutions made in this mutant resulted in a folding defect and hence an unstable enzyme.

The results show that Asp443 is vital for the RNase H activity of HIV-1 RT and that the Asp498 residue plays an important role in directing the correct folding of the protein.

DECLARATION.

I hereby declare that this dissertation is my own unaided work and that it has not been submitted for a degree at any other university.

R Brooksbank. 16.6.93

Richard Leslie Brooksbank (BSc Hons)

To my wife, Heather and my late grandfather.

ACKNOWLEDGEMENTS.

I wish to thank the South African Institute for Medical Research, the University of the Witwatersrand and the Foundation for Research Development for financial support during the course of this project. I thank Professor J. van den Ende for providing the facilities for the studies in this dissertation. I gratefully acknowledge the help, support and guidance of my supervisor, Professor Valerie Mizrahi.

Many thanks to the members of the Molecular Biology Unit of the S.A.I.M.R, Lindsay Dudding and Cyril Nkabinde, for their help, advice and companionship. In addition I would like to thank the members of my family for their help and encouragement, in particular my sister Judy for aiding me in the compilation of this manuscript and my father and mother for taking an interest beyond the call of duty.

Finally, I would like to thank my wife, Heather, for her understanding, patience and just for being there.

TABLE OF CONTENTS.

Preliminaries.		
Title page		i
Abstract		ii
Declaration		iii
Dedication		iv
Acknowledgements		v
1.	**Introduction.**	
1.1	Background.	1
1.2	The structure of HIV-1.	3
1.3	The life-cycle of HIV-1.	5
1.3.1	Entry of HIV-1 into the host cell.	5
1.3.2	Synthesis and integration of proviral DNA.	7
1.3.3	Transcription of viral genes.	7
1.3.4	The organisation of the HIV-1 genome.	8
1.3.5	Assembly and release of the mature virus.	11
1.4	HIV-1 RT.	12
1.4.1	The structure of HIV-1 RT.	13
1.4.1.1	The domain structure of HIV-1 RT.	13
1.4.1.2	The three-dimensional structure of HIV-1 RNase H.	16
1.4.1.3	The three-dimensional structure of HIV-1 RT heterodimer.	22
1.4.2	The function of HIV-1 RT.	27

1.4.2.1	The RNA- and DNA-dependent DNA polymerase activity of HIV-1 RT.	28
1.4.2.2	The RNase H activity of HIV-1 RT	30
1.4.2.3	The spatial and temporal relationship of the polymerase and RNase H activities.	34
1.4.3	The process of reverse transcription.	36
1.4.3.1	Initiation of reverse transcription.	37
1.4.3.2	The synthesis of (-)-strand DNA.	40
1.4.3.3	The synthesis of (+)-strand DNA.	42
1.4.3.4	The mechanism of DNA strand transfer during reverse transcription.	44
1.4.4	Mutagenesis studies of HIV-1 RT.	46
1.4.4.1	Mutagenesis studies of the active site of the polymerase domain of HIV-1 RT.	49
1.4.4.2	Mutagenesis studies of the active site or the RNase H domain of HIV-1 RT.	51
1.5	Aims	57
2.	Materials and Methods.	
2.1	Materials.	60
2.1.1	Reagents.	60
2.1.2	Oligonucleotides	61
2.1.3	Bacterial strains and plasmids.	62
2.2	Methods.	62
2.2.1	Oligonucleotide-directed mutagenesis.	62
2.2.2	5'-[^{32}P]-end labelling of oligonucleotides.	64
2.2.3	Screening of mutant M13-RT clones.	65
2.2.4	Construction of the mutagenised HIV-1 RT expression system.	67
2.2.4.1	Excision of the mutagenised *Asp*718 fragment.	67

2.2.4.2	Electroelution.	70
2.2.4.3	Reconstruction of M13-RT^m.	70
2.2.4.4	Rescreening of plaques.	71
2.2.4.5	Construction of the AR120/pDPTPRO4+pGALKRTE expression system.	72
2.2.5	Expression of authentic RT enzyme.	74
2.2.5.1	Small-scale induction of HIV-1 RT expression in AR120/pDPTPRO4+pGALKRTE.	74
2.2.5.2	Small-scale RT activity screening of mutants.	75
2.2.5.3	Large-scale induction of HIV-1 RT expression in AR120/pDPTPRO4+pGALKRTE.	78
2.2.6	Enzyme purification.	79
2.2.6.1	Cell lysis.	79
2.2.6.2	Phosphocellulose P11 column chromatography.	79
2.2.6.3	Q-Sepharose column chromatography.	81
2.2.6.4	Sephadex G-75 column chromatography.	82
2.2.7	RNase H assays.	83
2.2.7.1	Synthesis of the (+)-GAG^{345} RNA transcript.	83
2.2.7.2	Purification of the (+)-GAG^{345} RNA transcript.	84
2.2.7.3	Gel electrophoretic RNase H assay.	85
2.2.8	Steady-state kinetic studies of the D443A and wild-type enzymes.	86
2.2.9	Thermal inactivation studies of the D443A and wild-type RT enzymes.	86
3.	Results.	
3.1	Construction of mutants.	88
3.2	Analysis of mutants.	99
3.2.1	Protein purification.	99

3.2.2	Steady-state kinetic analysis of the RT activity of the D443A mutant.	108
3.2.3	Thermal inactivation studies of the D443A and wild-type RT enzymes.	108
3.2.4	Effect of the D443A mutation on the RNase H activity.	110
4.	Discussion and conclusion.	114
5.	Appendices.	128
6.	References.	134

1.) INTRODUCTION.

1.1) Background

The Human Immunodeficiency Virus Type-1 (HIV-1) is the causative agent of the Acquired Immunodeficiency Syndrome (AIDS). AIDS currently presents a major health threat throughout the world and unprecedented attention has been focused on HIV-1 by the scientific community in recent years.

HIV-1 is classified as a member of the retrovirus family which are defined as RNA-containing viruses that replicate through a DNA intermediate by means of a virally encoded RNA-dependent DNA polymerase, also called reverse transcriptase (RT) (Weiss *et al.*, 1982). The virus is further defined as a lentivirus or "slow" virus, a sub-family of the retroviruses, on the basis of genetic, pathological and morphological criteria. This group includes the equine infectious anaemia virus, the visna-maedi virus and the immunodeficiency viruses.

The first retrovirus to be identified in mammals, the murine leukaemia virus, was isolated by Gross in 1951 from an inbred mouse strain (Gross, 1951). Subsequently large numbers of cancer causing retroviruses were identified in species ranging from chickens to non-human primates. However, it was not until 1980 that the first human retrovirus was isolated, this being the human T-cell leukaemia virus type-1 (HTLV-I) isolated by Poiesz in the laboratory of Gallo and associates (Poiesz *et al.*, 1980).

This discovery was closely followed by the discovery of a second human retrovirus isolated from a patient with a T-cell variant of hairy cell leukaemia (Kalyanaraman *et al.*, 1982). This virus was subsequently named HTLV-II.

The isolation of HIV-1, the third human retrovirus, was the direct result of the search for the causative agent of AIDS. AIDS was first identified in 1981 after an outbreak of a range of conditions associated with immunodeficiency was found among a group of homosexual men in San Francisco (Centers for Disease Control, 1981). The syndrome soon spread to other risk groups, including intravenous drug abusers and haemophiliacs, and the number of recorded cases of the syndrome rapidly increased in the United States during the early 1980's. The etiology of the syndrome strongly suggested the involvement of a transmissible agent spread through genital secretions and blood.

The etiological agent of AIDS was eventually identified by Gallo *et al.* in 1984 by long-term propagation of viruses from several AIDS patients using permanent CD4+ T-cell lines (Popovic *et al.*, 1984; Gallo *et al.*, 1984). The availability of continuous, high-titre producer cell lines led to the development of highly purified viral reagents which provided the necessary material for the characterisation of the virus and the serological identification of exposed individuals (Sarngadharan *et al.*, 1984; Schupbach *et al.*, 1984). The virus was originally described as HTLV-III by Gallo because of the tropism of the virus. Similar viruses however were isolated by Levy and

designated as AIDS-related viruses (ARV) (Levy *et al.*, 1984). Eventually HTLV-III and ARV were found to be variants of the same AIDS virus and in 1986 an international committee recommended the name human immunodeficiency virus (HIV) for all isolates of the AIDS virus.

1.2) The Structure of HIV-1.

The viral particle of HIV-1, as revealed by electron microscopy, is approximately 110 nm in diameter and consists of a central cylindrical core surrounded by an outer envelope (see Figure 1). The core of the virion is composed of p24 as a major capsid protein and envelopes the two copies of 35S single-stranded viral ribonucleic acid (RNA). This RNA is found in the form of a ribonucleoprotein complex with the viral RT (p66/p51), integrase (p32), protease (p11) and the RNA-binding nucleocapsid protein (p7). The myristilated Gag protein (p17) forms the outer shell of the virion with its amino-terminal region inserted into the lipid membrane envelope which is derived from the cell membrane of the infected host cell. The extracellular and transmembrane glycoproteins gp120 and gp41 exist in the form of a non-covalent complex on the outer surface of the envelope.

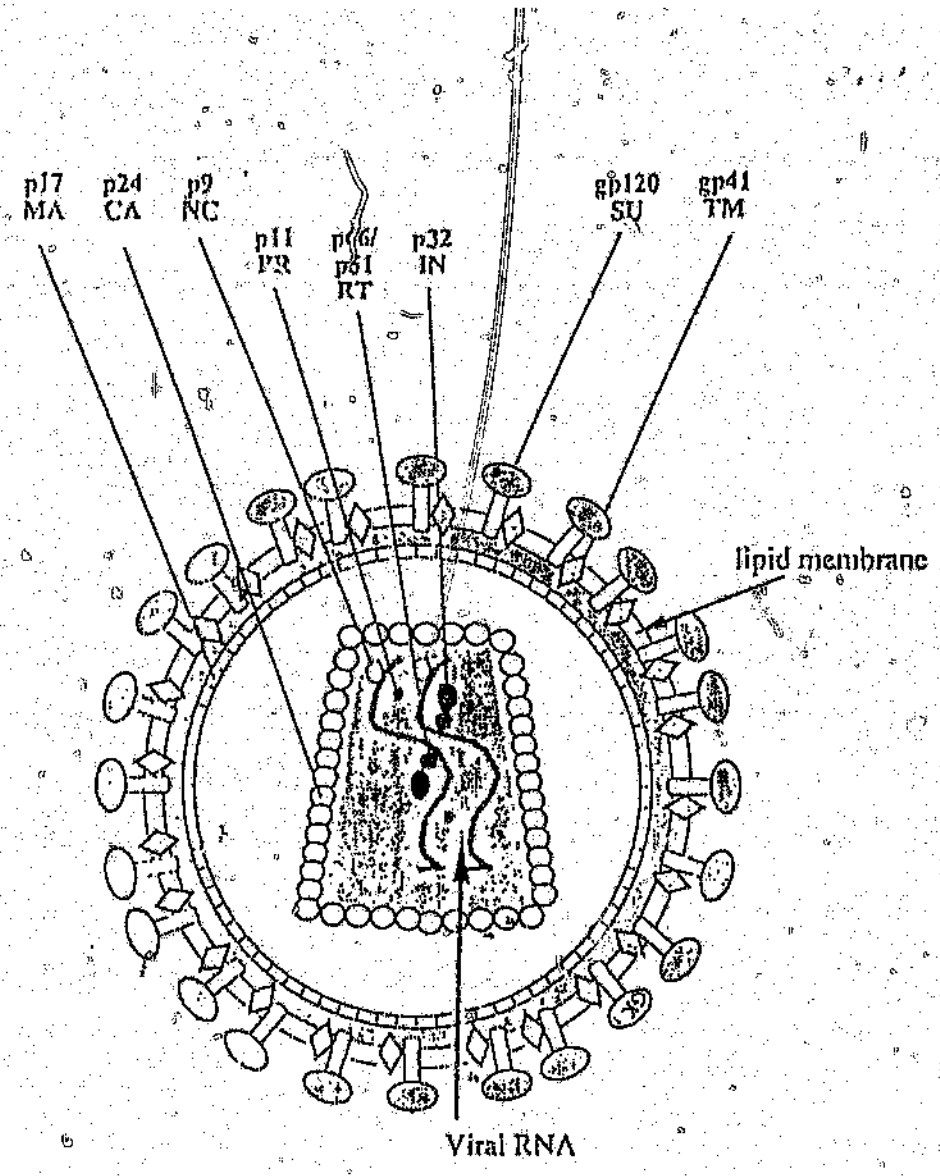


Figure 1: The Structure of HIV-1. (Adapted from Vaishnav Wong-Staal, 1991)

1.3) The Life-cycle of HIV-1.

The life cycle of HIV-1 can be separated into four main stages (see Figure 2): Infection of the host cell by the virus, reverse transcription and integration of the viral genome into the DNA of the infected cell, transcription and translation of the viral genes, and finally, reassembly and budding of the virus.

1.3.1) Entry of HIV-1 into the Host Cell.

In most cell types, entry of HIV-1 is initiated by an interaction between the viral envelope glycoprotein gp120 and the CD4 receptor molecule present on the surface of the target cell. Although CD4 appears to be the primary receptor molecule for HIV-1, especially in relation to the infection of helper T-cells, it is thought that additional factors may be required for the initial binding of the virus to other cell types where CD4 has been shown to be absent from the cell surface (Clapham *et al.*, 1989). Following binding of the virus to the cell surface, fusion of the viral envelope to the host cell membrane is mediated by the transmembrane envelope glycoprotein gp41 (Kowalski *et al.*,1987; Gallaher, 1987) in a pH-independent manner (Stein *et al.*, 1987) possibly by internalisation of the CD4 receptor after binding with HIV-1 gp120.

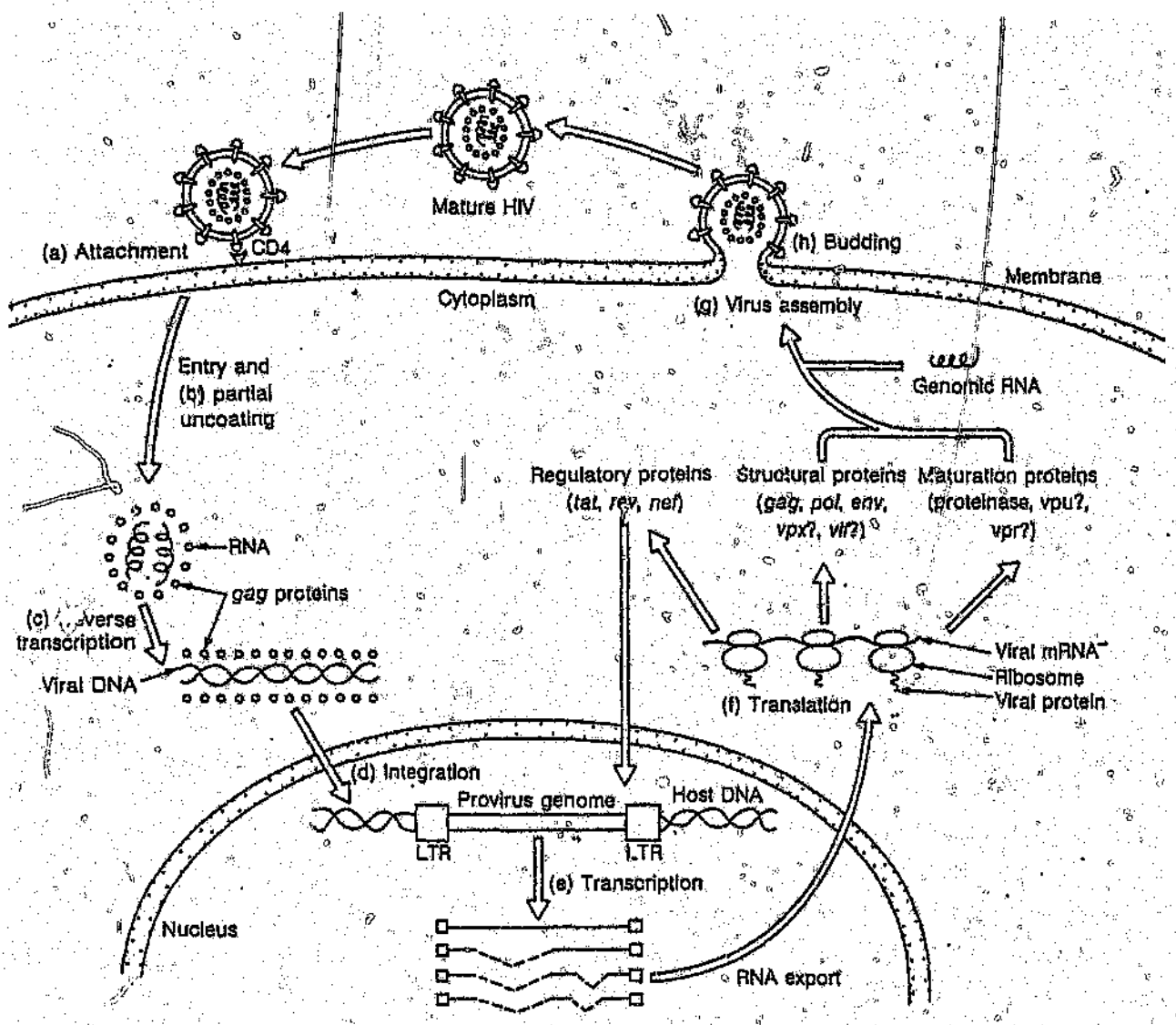


Figure 2: The Life-cycle of HIV-1. (Adapted from Cann and Karn, 1989)

1.3.2) Synthesis and Integration of Proviral DNA.

Reverse transcription of the diploid viral RNA genome into a double-stranded DNA molecule occurs soon after the viral core particle has entered the cytoplasm of the host cell. This process is described in greater detail in section 1.4.3. The integration of the proviral double-stranded DNA into the host genome is mediated by the viral integrase protein, a 31-kD product of the *pol* gene. Most aspects of integration of HIV-1 DNA are similar to other retroviral systems in that the integrated DNA contains highly conserved di-nucleotide sequences (5' TG---CA 3'), and there is duplication of a short stretch of cellular DNA at the site of integration. In addition HIV-1 termini contain short inverted repeats which may be important for specificity of integrative recombination with respect to the viral DNA (Vaishnav and Wong-Staal, 1991).

1.3.3) Transcription of Viral Genes.

The gene expression of HIV-1 is controlled by an interplay between *trans*-acting cellular transcription factors, *cis*-acting regulatory sequences present in the viral long terminal repeat (LTR) and the viral *trans*-activators Tat and Rev. The regulatory sequences within the LTR of HIV-1, upstream of the RNA initiation site, can be broadly divided into three regions: the promoter, enhancer and negative regulatory regions. These are all sequences typically involved in transcription by cellular RNA polymerase II and include the TATA box promoter and recognition sequences

for the T-cell specific transcription factor NFκB (Nabel and Baltimore, 1987) and the ubiquitous transcription factor Sp1 (Jones *et al.*, 1986). In addition to these upstream regulatory sequences the HIV-1 promoter contains several protein binding sites downstream of the RNA start site as well as a potential intragenic enhancer sequence in the *gag-pol* region of the viral genome.

The exact role of these sequences and their corresponding factors in basal or induced transcription from the HIV-1 promoter is however still unclear.

1.3.4) The Organisation of the HIV-1 Genome.

The HIV-1 genome consists of the typical retroviral genes *gag*, *pol*, and *env* as well as the regulatory genes *tat*, *rev* and *nef* and the accessory genes *vif*, *vpr*, *vpu*, *vpt* and *tev/tnv* (see Figure 3).

The *gag* gene of HIV-1 is initially translated from the full length viral messenger RNA (mRNA) as a polyprotein precursor, $Pr55^{gag}$ with a domain structure of NH_2-p17(MA)-p24(CA)-p9(NC)-p7-COOH. The precursor is processed by the viral protease (Mervis *et al.*, 1988) in order to give rise to individual proteins some of which play important roles in the assembly and release of the mature virus particle (see section 1.3.4 and 1.3.5).

The *pol* gene encodes for the catalytic proteins of HIV-1

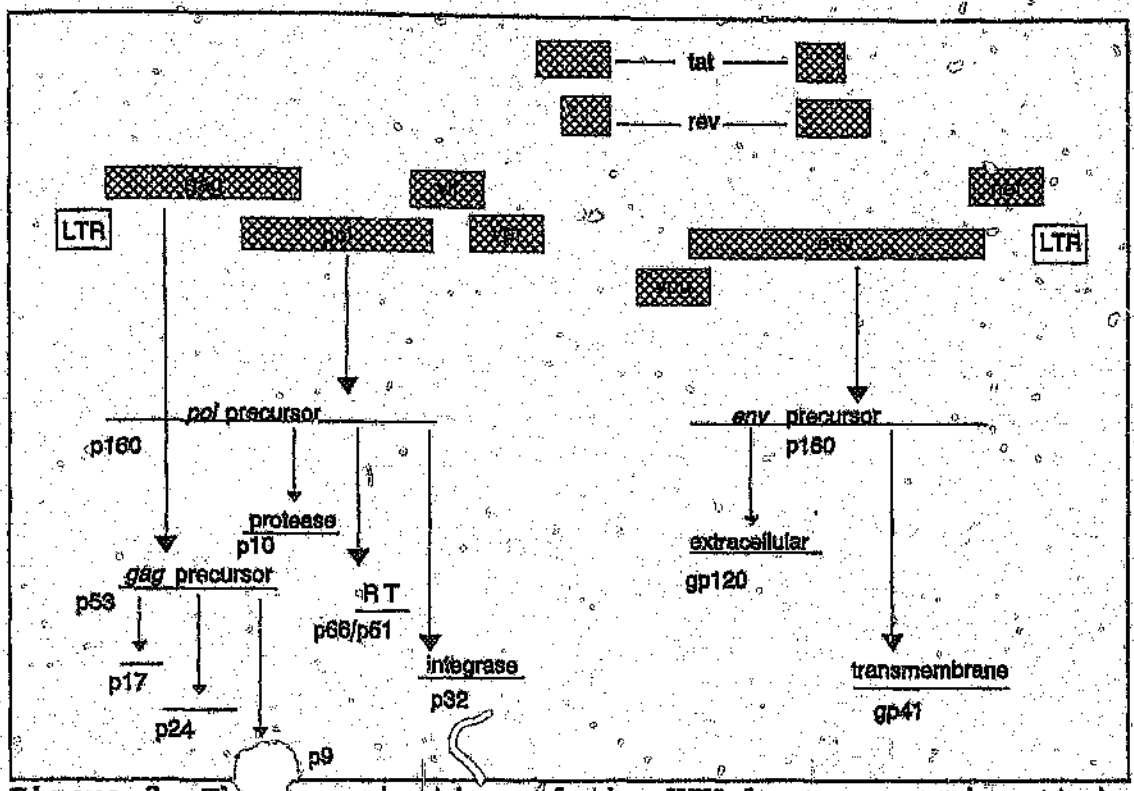


Figure 3 The organisation of the HIV-1 genome and protein processing pathways. (adapted from Rosen, 1991)

namely the protease, integrase and the RT. As with other retroviruses it is translated from the genomic length viral mRNA to produce the polyprotein precursor $Pr160^{gag\text{-}pol}$ which has a domain structure of NH_2-p17(MA)-p24(CA)-p9(NC)-p7(NC)-p10(PR)-p66(RT)-p32(INT)-COOH. The *gag* and *pol* genes overlap by 241 nucleotides with the *pol* gene being in a -1 reading frame relative to the *gag* reading frame. Translation of the *pol* gene is brought about by a ribosomal frame shift which occurs at a low frequency of approximately 5% and is directed by a short homopolymeric sequence at the *gag-pol* overlap region which is followed by a hairpin loop structure (Coffin, 1990). This strategy of gene expression ensures that the structural proteins derived from the *gag* gene are produced in larger quantities than

the enzymatic products of the *pol* gene.

As for the *gag* gene products, the *pol* precursor is cleaved by the viral protease to produce the p32 integrase, the p10 protease and the p66/p51 RT proteins. The production of the RT p66/p51 heterodimer and the role of the subunits in RT activity is discussed in greater detail in section 1.4.1.

The envelope glycoproteins of HIV-1 are initially synthesized as an 88-kDa protein inserted into the rough endoplasmic reticulum, where the addition of high-mannose N-linked carbohydrate chains as well as folding into an appropriate tertiary structure takes place (Fennie and Lasky, 1989). The carbohydrate chains are terminally modified in the golgi complex. The resulting gp160 precursor is cleaved in a nonlysosomal acidic compartment by a cellular protease, and the mature envelope proteins (gp120 and gp41) are transported to the cell surface and inserted into the cell membrane prior to viral assembly.

The regulatory proteins Tat, Rev and Nef and the accessory protein Tev/Tnv are all translated from singly or multiply spliced small mRNAs that are produced shortly after integration of the provirus into the host genome and hence are known as the early mRNAs. The remainder of the accessory proteins Vif, Vpu, Vpt and Vpr are translated from intermediate length transcripts produced at a later stage along with the full length mRNA of the *gag* and *gag-pol* genes.

1.3.5) Assembly and Release of the Mature Virus.

The selective packaging of viral genomic RNA and the assembly of the mature virus particle is initiated by the aggregation of the polyprotein precursors $Pr55^{gag}$ and $Pr160^{gag-pol}$ under the plasma membrane, a process mediated by the *gag* derived protein p17. The precursor molecules of HIV-1, in common with other retroviruses, are posttranslationally modified at the amino terminus by the addition of myristic acid which facilitates the hydrophobic interaction of the precursors with the plasma membrane (Mervis *et al.*, 1988; Veronese *et al.*, 1988). Abolition of myristylation by site-directed mutagenesis of the amino-terminal glycine residue of the Gag precursor was shown to cause a complete disruption of the assembly process (Gottlinger *et al.*, 1989).

The viral genomic RNA is brought to the site of assembly by the p9 nucleocapsid domain of the Gag and Gag/Pol precursors. The interaction of p9 with the genomic RNA is determined by features of both the RNA and the protein. HIV-1 genomic RNA has been shown to contain a major packaging signal sequence located between the U5 LTR and the *gag* initiation codon which is essential for efficient viral RNA packaging (Lever *et al.*, 1989). Removal of this sequence has been shown to result in a 98% reduction in efficiency. The p9 nucleocapsid protein, in common with other retroviral nucleocapsid proteins, contains copies of the sequence Cys-X2-Cys-X4-His-X4-Cys which is similar to the "zinc finger" motif associated with many DNA-binding proteins

(Berg, 1986). Mutation or deletion of this sequence in p9 results in a marked decrease in the genomic RNA content of mature virions. The genomic RNA is encapsulated by the major core protein p24 Gag which forms a protein shell around the nucleocapsid.

The Gag and Gag/Pol precursors of HIV-1 are cleaved by the viral protease during or after budding to produce the individual proteins. This process is responsible for the morphological maturation of the virion in which the core condenses to form an electron dense cylindrical structure (Gottlinger *et al.*, 1989). The fact that processing of the precursors occurs at a late stage of maturation suggests that the viral protease is only weakly active in the precursor form and requires a high local concentration of precursor for sufficient activity to occur. Alternatively the protease may be active only as a dimer and hence protease activity may only occur once the precursors have correctly aligned themselves in the later stages of budding. Investigations of the structure of HIV-1 protease have shown that the enzyme can form dimers with the active site being formed by the dimer linkage region (Skalka, 1989).

1.4) HIV-1 RT.

HIV-1 RT is the enzyme solely responsible for the conversion of the single-stranded RNA genome of HIV-1 into double-stranded proviral DNA prior to integration of the proviral DNA into the host cell genome. The enzyme therefore plays an

essential role in the life cycle of HIV-1 and hence has become a major target of research in an attempt to find effective therapeutic agents against AIDS.

RTs of viral origin were first identified in 1970 (Baltimore, 1970). The purification to homogeneity of the RT from the avian myeloblastosis virus by Kacian *et al.* (1971) and Moelling *et al.* (1971) led to the identification of three enzymatic activities associated with viral reverse trancriptases: a RNA-dependent DNA polymerase activity, a DNA-dependent DNA polymerase activity and a ribonuclease H (RNase H) activity (Goff, 1990) which removes the RNA moiety from a DNA.RNA hybrid. However, although a large number of RTs were identified during the course of the 1970's, the enzyme only became a major research target with the onset of the HIV-1 epidemic in the early 1980's.

1.4.1) Structure of HIV-1 RT.

1.4.1.1) The Domain Structure of HIV-1 RT.

As discussed in section 1.3, HIV-1 RT is initially cleaved by the viral protease from the *gag-pol* precursor to give rise to a 66-kDa protein consisting of 590 amino acids. In virions the enzyme exists as a p66/p51 heterodimer (Di Marzo Veronese *et al.*, 1986) which is formed by a second cleavage event whereby a 15-kDa carboxy-terminal segment is removed from the p66 protein to give rise to a 51-kDa truncated version of p66. Whether cleavage of p66 occurs *in vivo* when the protein is found as a monomer or as

a homodimer (p66/p66) is unclear at present. However there is evidence to suggest that cleavage occurs in the homodimer form with the first cleavage event causing a steric change that prevents cleavage of the remaining p66 subunit (Davies *et al.*, 1991; Kohlstaedt *et al.*, 1992). This will be discussed further in section 1.4.2 and 3. The resulting heterodimer therefore consists of two subunits, the p66 and p51 subunits, with identical amino-termini. Comparative mapping studies of HIV-1 RT with other retroviral RTs have shown that the polymerisation activity maps to the amino-terminal portion of the p66 chain and that the RNase H activity corresponds to the carboxy terminal 15-kDa segment that is removed during proteolytic processing (Johnson *et al.*, 1986). The p51 subunit therefore lacks the RNase H domain whilst still possessing the entire polymerase domain. RTs of other retroviruses have been characterised as both monomers and dimers. The RT of murine leukaemia virus (MuLV) is found as an 80-kDa monomer and is typical of RTs found in the murine and feline leukaemia virus families (Goff and Lobel, 1987). AMV RT however is a heterodimer with α and β subunits of 95- and 63-kDa respectively with both subunits showing polymerase and RNase H activity (Goff, 1990). The fact that the p66 subunit of HIV-1 RT appears to contain all of the sequence information required for RT activity has led to an extensive investigation of the role played by the p51 subunit in HIV-1 RT. Reports have shown that the purified recombinant homodimer, p51/p51 was either inactive or only partially active relative to the p66/p51 heterodimer (Hansen *et al.*, 1988; Hizi *et al.*, 1988; Starnes *et al.*, 1987). Le Grice *et al.* (1991) showed that wild type p51

failed to restore polymerase activity to a heterodimer whose p66 subunit was mutated at the proposed active site, whereas the wild type p66 formed an active heterodimer with an identically mutated p51 subunit. However other independent groups have reported appreciable polymerase activity attributable to the p51 subunit using i) activated gel analysis of the viral enzyme (Lori *et al.*, 1988), ii) a recombinant p51 with a modified COOH-terminus (Restle *et al.*, 1990). Since the p51 protein lacks the RNase H domain, no RNase H activity has been attributed to this subunit. The p66/p66 homodimer produced in bacterial mutants lacking proteases is a loosely associated dimer with a specific activity less than half that of the heterodimer. Investigations into the active site of the polymerase domain by competition studies using different primer-template complexes have identified [illegible] binding site in the heterodimer (Dirani-Diab *et al.*, 1992) and binding of non-nucleoside inhibitors to HIV-1 RT resulted in selective binding to the p66 chain in the heterodimer (Wu *et al.*, 1991).

The region important for heterodimer formation was shown by Becerra *et al.* (1991) to reside in a central region of the p66 subunit. They also inferred that the carboxy-terminal region corresponding to the RNase H domain influenced efficient formation of the heterodimer. Although it appears clear at present that the p66 subunit is the major catalytic component of HIV-1 RT and that the p66/p51 heterodimer is the most active and stable form of the enzyme, the contribution made by p51 to RT activity is still speculative.

Although comparative mapping studies have clearly identified the functional domains of HIV-1 RT, the two domains have been shown to possess a large degree of structural interdependence. For some retroviral RTs, in particular MuLV RT, it is possible to physically separate the polymerase activity from the RNase H activity of the enzyme with no loss of specific activity in either case (Tanese and Goff, 1988). However, in the case of HIV-1 RT, mutations within the RNase H region of the enzyme can reduce or abolish the polymerase activity, and likewise, mutations within the polymerase domain can disrupt RNase H activity (Mizrahi *et al.*, 1990; Schatz *et al.*, 1989; Hizi *et al.*, 1990; Prasad and Goff, 1989). Hostomsky *et al.* (1991) demonstrated that although the p15 RNase H domain had no detectable RNase H activity when expressed *in vitro* on its own, RNase H activity was reconstituted by the addition of the purified p51 subunit to the protein.

1.4.1.2) The Three-Dimensional Structure of HIV-1 RNase H.

The first three dimensional structure of a RNase H enzyme to be published was that of the *E.coli* RNase HI by Katayanagi *et al.* (1990). The structural gene of *E.coli* RNase HI, *rnhA*, encodes a monomer polypeptide of 155 amino acid residues. Although the *E.coli* enzyme shows a relatively low homology at the amino acid level with the HIV-1 RNase H domain, a number of key amino acids are shared. The structure of the RNase H domain of HIV-1 RT was published by Davies *et al.* (1991). The RNase H construct used for the study was that of Hostomsky *et al.* (1991) and consisted of

the RNase H domain with an additional fourteen amino-terminal amino acids. This stretch of additional amino acids was upstream of the usual viral protease cleavage site between Phe440 and Tyr441 that defines the carboxy-terminus of the p51 subunit (Mizrahi *et al.*, 1989A; Le Grice *et al.*, 1989; Becerra *et al.*, 1990; Graves *et al.*, 1990). The addition of the short stretch of amino acids gave rise to a structurally stable RNase H domain that could be crystallised, despite its lack of enzymatic activity.

Comparison of the structure of *E.coli* RNase HI with that of the HIV-1 RNase H domain revealed that the two proteins show strong similarities in their overall structure, however there are some important differences (see Figure 4). Both proteins adopt an α-ß form, the RNase H domain of HIV-1 being folded into a five stranded ß sheet flanked by an asymmetric distribution of four α helices. The *E.coli* RNase HI consists of five ß sheets surrounded by an asymmetric distribution of five α helices. The most important feature shared by the two structures is the area of divalent metal ion binding at the proposed catalytic site of the enzyme. The presence of divalent metal ions has been shown to be essential for RNase H activity (Starnes and Cheng, 1989). It is in this region that HIV-1 RNase H and *E.coli* RNase HI share the greatest degree of homology at an amino acid sequence level. The four conserved acidic amino acids found at this site, Asp10, Glu48, Asp70 and Asp134 in *E.coli* versus Asp443, Glu478, Asp498 and Asp549 in HIV-1 RNase H, are among seven amino acids conserved among all known retroviral and bacterial RNase H

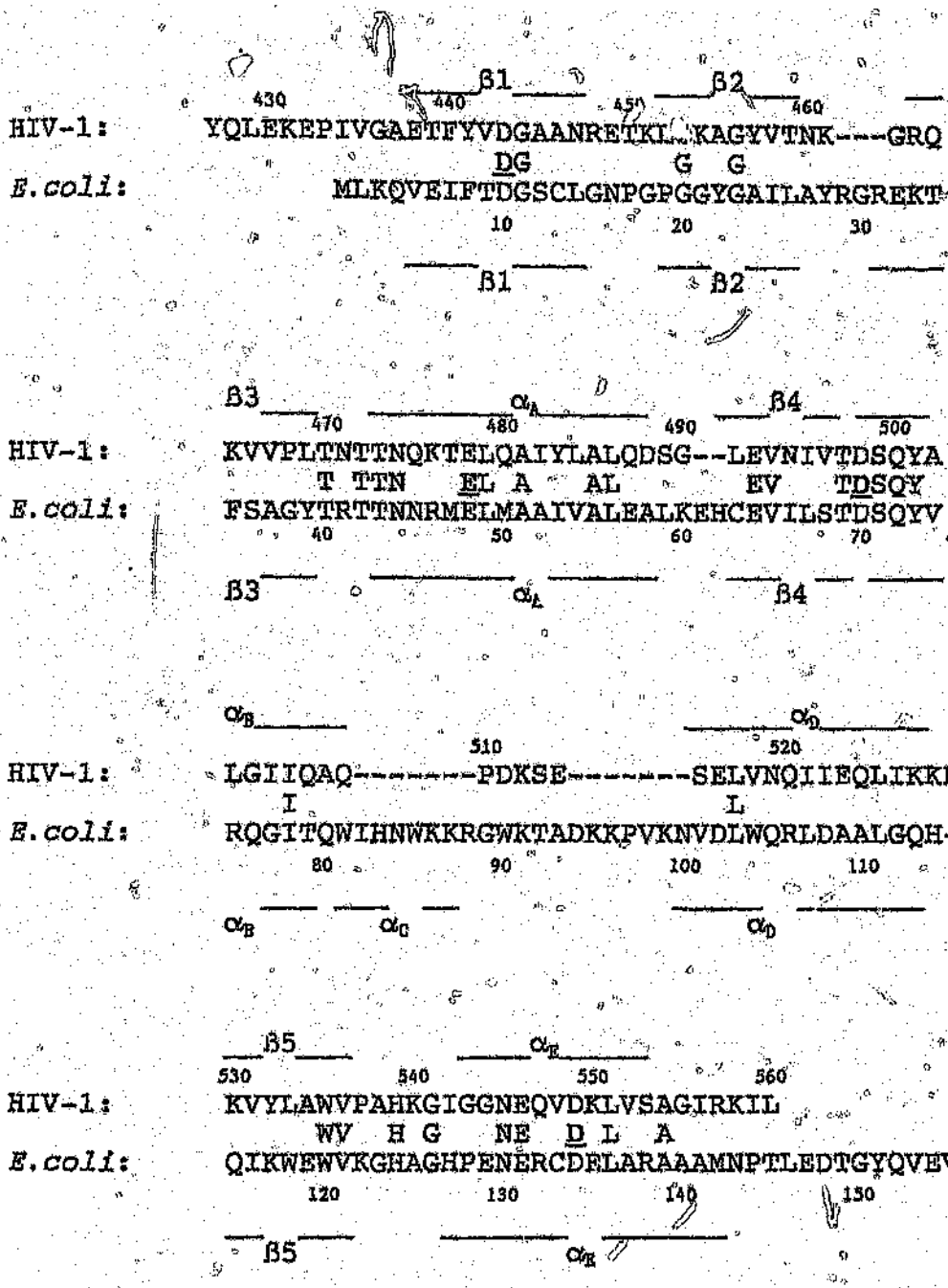


Figure 4: Amino acid sequence and geometrical alignment of secondary structural elements in the RNase H domain of HIV-1 RT and in *E.coli* RNase HI. Identical residues in the structures are indicated. The four catalytically important acidic residues are shown in bold and underlined. (Adapted from Davies *et al.*, 1991)

sequences (Doolittle *et al.*, 1989). These residues are found within the five strand ß sheet core of both proteins and, in the HIV-1 RNase H domain, were reported to co-ordinate two divalent metal ions whereas in *E.coli*, co-ordination of only one metal ion was reported. This apparent discrepancy may have arisen as a result of a crystal packing effect in the case of the *E.coli* enzyme (Davies *et al.*, 1991) or may be an artefact caused by the high concentrations of Mn^{2+} solution used for the study of the sites of divalent metal ion binding in the crystal structure of the HIV-1 RNase H enzyme (Katayanagi *et al.*, 1992). Asp443, Asp498 and Glu478 and the corresponding residues in *E.coli* have all been shown to be essential for RNase H activity in both enzymes (Mizrahi *et al.*, 1990; Schatz *et al.*, 1989; Kanaya *et al.*, 1990). The roles of these residues in RNase H activity will be discussed in greater detail in section 1.4.4.2.

Another noteworthy feature of the structure of HIV-1 RNase H is the location of the viral protease cleavage site used for the formation of the heterodimer. As stated earlier (section 1.4.1), the HIV-1 protease cleaves the p66 protein between Phe440 and Tyr441 and in only one of the p66 subunits of the p66/p66 homodimer. The three dimensional structure of HIV-1 RNase H reveals that the protease cleavage site in the properly folded protein is located at position three and four of the ß1 strand which is a long strand forming the centre of the beta sheet. In this position it would appear to be physically impossible for the viral protease to cleave this site in the properly folded protein. This observation has two important consequences for HIV-

1 RT. Davies *et al.* (1991) have proposed that since the p66/p66 homodimer possesses RNase H activity similar to that of the heterodimer (Schatz *et al.*, 1989) at least one of the RNase H domains must be correctly folded within the homodimer structure. This implies that the p66 homodimer is asymmetric with the other RNase H domain being unfolded to an extent that allows the viral protease to gain access to the cleavage site. The net consequence of this is that only one of the two p66 proteins are cleaved hence giving rise to the formation of the p66/p51 heterodimer and explaining the 1:1 ratio of p66 to p51 found *in vitro* and *in vivo*. Davies *et al.* (1991) also proposed that the RNase H domain plays a structural role in stabilising the heterodimer by interacting with the polymerase domains. The importance of the carboxy-terminal region of HIV-1 RT in dimerisation and stability has been confirmed by Becerra *et al.* (1991). The second consequence of cleavage at Phe440-Tyr441 is that the carboxy-terminal 120 residues that are released after cleavage are unlikely to be able to form a properly folded active fragment since the first three residues of the central β-sheet are lost. This would explain why the normal cleavage product of heterodimer formation shows a 1000-fold lower RNase H activity than the intact p66 precursor peptide. The position of the cleavage site also explains the necessity for the addition of fourteen amino acids to the amino-terminus of the RNase H domain to give rise to a stable and crystallisable RNase H domain (Hostomsky *et al.*, 1991). The instability of the cleavage product would also make the polypeptide susceptible to further degradation by proteases, hence the lack of success experienced by groups attempting to

detect this fragment during proteolytic processing of RT.

Two features of the structure of *E.coli* RNase HI stand out as being significantly different from that of the HIV-1 RNase H domain. The first is the presence of an additional α helix (α_C) and an extended loop between α_C and α_D that are absent from HIV-1 RNase H. This is a lysine and tryptophan rich region that has been implicated in the binding of the DNA.RNA hybrid prior to digestion of the RNA moiety. The absence of this structure in the HIV-1 RNase H domain may explain the lack of RNase H activity and would imply that the polymerase domain is necessary for the binding and correct alignment of the DNA.RNA substrate. The RNase H domain of MuLV RT possesses approximately the same number of amino acid residues in this region as *E.coli* RNase HI and this may account for its ability to act as an RNase H independent of the polymerase domain.

The second feature of *E.coli* RNase HI that is markedly different from that of HIV-1 RNase H is the loop that connects ß5 to α_E and passes across the front of the ß-sheet (see Figure 4). In the RNase H of HIV-1 RT this loop, which corresponds to residues 538 to 542, was disordered (Davies *et al.*, 1991). The loop contains the universally conserved residue His539. Mutations of this residue in *E.coli* RNase HI have been shown to affect catalytic efficiency and substrate binding (Kanaya *et al.*, 1990) whereas similar mutations in HIV-1 RNase H decrease both polymerase and RNase H activities (Schatz *et al.*, 1989). The disorder of this loop in the HIV-1 RNase H structure may account

for the loss RNase H activity and it is possible that the polymerase domain may stabilise the structure.

1.4.1.3 The Three Dimensional Structure of the HIV-1 RT Heterodimer.

Attempts to determine the three dimensional structure of the HIV-1 RT heterodimer at high resolution have only recently met with success after several years of frustrating difficulties. It was originally suggested that the inability of HIV-1 RT to form sufficiently ordered crystals for high resolution structural studies may have resulted from the inherent flexibility of the enzyme. In order to circumvent this problem a number of strategies were employed to give rise to a more rigid protein structure that would be likely to form ordered crystals. Jacobo-Molina *et al.* (1991A) cocrystallised the antigen-binding fragments of monoclonal antibodies (Fab) with the RT heterodimer and obtained crystals that diffracted X-rays to a moderate resolution of 6 Å. The Fab fragment is thought to enhance crystallisation by acting as a molecular "clamp" and thereby restricting the conformational freedom of the RT. They were able to further enhance the resolution of these crystals to 3.5 Å by the addition of a 19-base DNA strand paired with an 18-base strand to form a 19/18 double-stranded DNA (dsDNA) oligomer with a single adenine base overhang at the 5′ end. The sequence of the 19-base strand corresponded to the primer binding site (pbs) of HIV-1. The complementary 18-base strand corresponded to the 3′-terminus of the human transfer RNA^{Lys3} ($tRNA^{Lys3}$) that is used as

a primer to initiate polymerisation by HIV-1 RT (see section 1.4.3.1). The 19/18 dsDNA oligomer therefore mimics the initial template-primer. The resulting RT/Fab/dsDNA complex may resemble HIV-1 RT during DNA-dependent DNA polymerisation.

Structural details of this complex at a resolution of 7-Å were published by Arnold *et al.* (1992). This group found that the dsDNA bound in a putative template-primer groove located on the surface of the enzyme. By using a mercurated derivative of UTP they were able to identify a potential site for the binding of nucleoside triphosphates which they suggested corresponded to the active site of polymerisation. This site was located close to the 5' terminus of the dsDNA at the position of the adenine overhang. The putative active site of the RNase H domain, identified as the site of Mg^{2+} binding, was found to face the dsDNA at a position approximately 15 base pairs downstream from the polymerase site. However the relatively poor resolution obtained in this study prevented a detailed description of the polypeptide backbone of the RT heterodimer.

In an independent report, Kohlstaedt *et al.* (1992) determined the crystal structure at 3.5 Å of the HIV-1 RT heterodimer complexed with an inhibitor of the enzyme, nevirapine (Merluzzi *et al.*, (1990)). The structure of the p66 subunit was found to be folded into five separate subdomains (see Figure 5A): the RNase H domain, whose coordinates matched those of the structure reported by Davies *et al.* (1991), and four subdomains

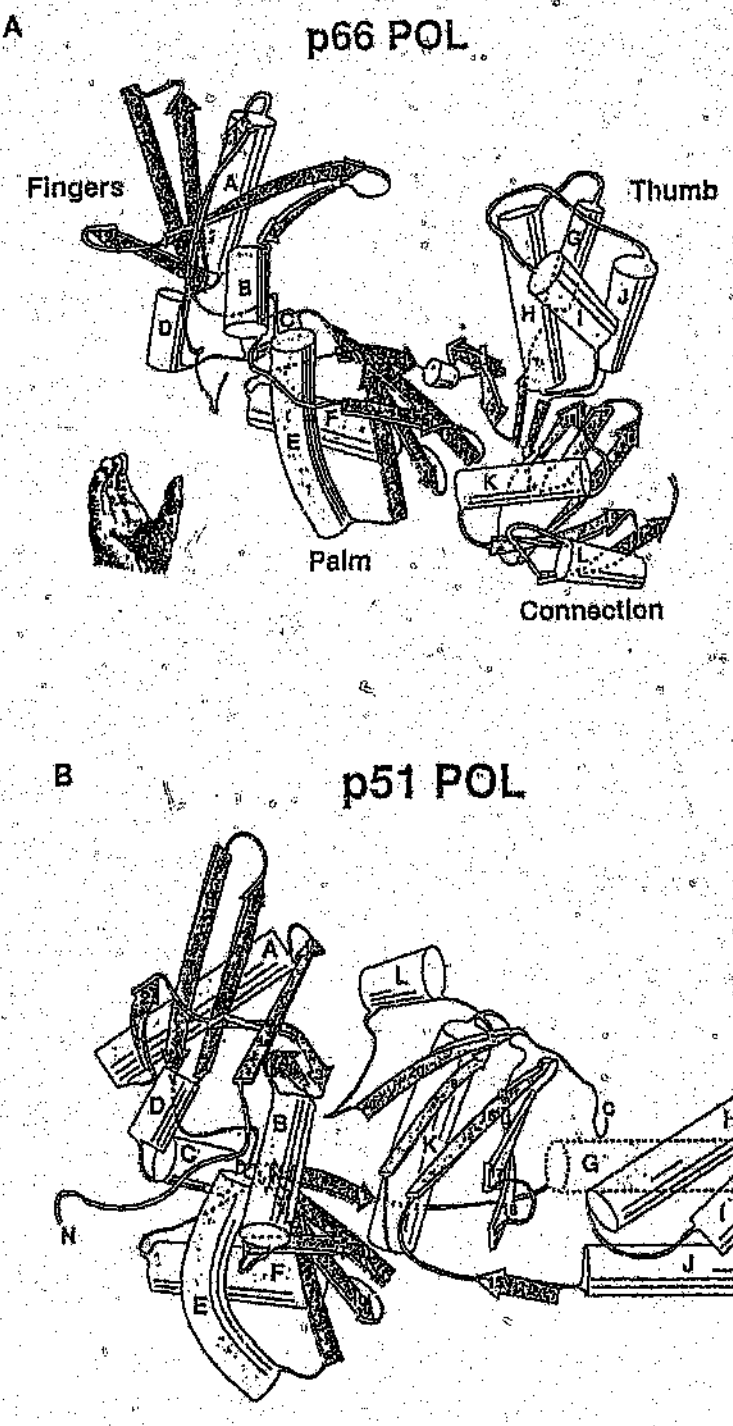


Figure 5: The structure of the p66 (A) and p51 (B) polymerase domains of HIV-1 RT. α helical regions are represented as tubes and ß strands as arrows. (Adapted from Kohlstaedt *et al.*, 1992).

of the polymerase region. The authors likened the structure of the polymerase region to that of a right hand and hence named the subdomains the "fingers", "palm" and "thumb". The fourth subdomain linked the hand domain of the polymerase region to the RNase H domain and was therefore called the "connection" subdomain. The gross feature of the p66 subunit that is most noticeable is the large cleft formed by the finger, palm and thumb subdomains of the polymerase region and the RNase H subdomain. The positioning of the finger, palm and thumb subdomains is analogous to the groove formed by the Klenow fragment (KF) of *E.coli* DNA polymerase I (Ollis *et al.*, 1985). However, although there are a few gross structural similarities between these two DNA polymerases, particularly in the palm and thumb subdomains, there is very little homology at a sequence level. This groove in RT was shown by the authors to be able to accommodate an A-form RNA.DNA hybrid with approximately 20 base pairs separating the polymerase an RNase H active sites. This RNA.DNA duplex binding groove in RT reported by Kohlstaedt *et al.* is the same as the template-primer groove described by Arnold *et al.* (1992).

The conformation of the p51 subunit is remarkably different from p66 given that the two proteins have the same amino acid sequence (see Figure 5B). The most notable difference between the two proteins is the lack of a groove in the p51 subunit. In p51, the groove is obscured by the connection subdomain and the thumb is rotated so as to be further from the fingers and palm thereby forming a more elongated structure compared to the p66 subunit.

The fingers and palm are also located closer together. The proposed polymerase active site is effectively buried which would explain why the p51 subunit is virtually devoid of polymerase activity (Hansen *et al.*, 1988).

The p51 subunit interacts with the p66 subunit in three places. The tip of the fingers and connection subdomains of p51 interact with the palm and connection subdomains of p66 while the extended thumb of p51 contacts the RNase H domain. These interactions contribute to the formation of the template-primer binding groove. The contact made by p51 with the RNase H domain of p66 stabilises the polypeptide loop containing His539 that was disordered in the structure of the RNase H fragment reported by Davies *et al.* (1991). This observation supports the proposal that disorder of this loop may account for the inactivity of the isolated RNase H fragment used by Davies *et al.* and that this loop may be stabilised either by the polymerase domain of the p66 subunit or by the p51 subunit. In addition it is evident from the structure that the substrate binding site is formed predominantly by the polymerase domain of p66 and the thumb region of p51 and that the RNase H domain on its own is unlikely to be able to bind the substrate correctly. Kohlstaedt *et al.* have proposed that the p51 subunit, based on its location relative to the A-form RNA.DNA duplex, could form part of the binding site for $tRNA^{lys3}$. In addition Barat *et al.* (1989) have shown that binding of $tRNA^{lys3}$ to both the p51 and p66 subunits occurs.

The marked asymmetrical structure of the HIV-1 RT

heterodimer supports the proposal of Davies *et al* (1990), that heterodimer formation occurs as a result of only one protease cleavage site being accessible to the viral protease when the enzyme is in a p66/p66 homodimer configuration. In order for the homodimer to form with the same points of contact that exist in the heterodimer, one of the p66 subunits must assume the structure of the p51 protein. This in turn would mean that the RNase H domain, prior to cleavage, would have to be in a partially unfolded state in order to be accommodated in the homodimer. This would expose the cleavage site and render it susceptible to the activity of HIV-1 protease thereby resulting in the formation of the p66/p51 heterodimer. The other p66 subunit would be in a stable configuration with the protease cleavage site buried, preventing further cleavage.

1.4.2) The Function of HIV-1 RT.

HIV-1 RT is a multifunctional enzyme and, as stated previously, possesses three main enzymatic activities that enable it to convert the single-stranded RNA viral genome into the double-stranded DNA proviral form prior to integration of the proviral DNA into the host genome. These are a DNA-dependent DNA polymerase activity, a RNA-dependent DNA polymerase activity and a RNase H activity (Goff, 1990). Recently a potential fourth activity, a ribonuclease D activity (RNase D), was also found to be associated with HIV-1 RT (Ben-Artzi *et al.*, 1992A). However this has subsequently been shown to result from contaminating *E.coli* RNase III that had copurified with the recombinant RT

(Ben-Artzi *et al.*, 1992B; Hostomsky *et al.*, 1992).

1.4.2.1) The RNA- and DNA-dependent DNA polymerase activity of HIV-1 RT.

A retrovirally associated RNA and DNA-dependent DNA polymerase or RT activity was first identified by Baltimore (1970) in the avian myeloblastosis virus. Further characterisation of RTs from a number of different systems provided information relating to the enzyme's substrate specificities (Baltimore and Smoler, 1971), kinetics and processivity (Fujinaga *et al.*, 1970; Watson *et al.*, 1979; Gregerson *et al.*, 1980), interactions with its cognate tRNA primer (Panet *et al.*, 1975) and fidelity of DNA synthesis (Skinner and Eperon, 1986).

With the initial characterisation of HIV-1 RT by Rho *et al.* (1981) basic properties of the enzyme were determined. It was found that the enzyme required a divalent metal ion for activity, preferably Mg^{2+}, and that the optimal conditions for activity were a pH of 7.5-8.0 and a salt concentration of between 50 and 100 mM (). The K_m of the enzyme for substrates was found to be 10 µM (Reardon and Miller, 1990) and the steady state K_m value of polymerisation on a poly(rA).oligo(dT) primer template was found to be 3.5 µM (Huber *et al.*, 1987).

More detailed investigations of HIV-1 RT in HIV-1 showed that synthesis of DNA is an ordered process in which template-

primer binding precedes binding of the nucleotide substrate and the processivity of the enzyme was found to be template-primer dependent (Huber *et al.*, 1989). Huber *et al.* (1989) also demonstrated that the enzyme was most processive on a poly(rA).oligo(dT) template, being capable of adding more than 300 nucleotides before dissociation of the enzyme from the template-primer. In comparison, incorporation of dTMP into poly(dA).oligo(dT) was found to be distributive with extension of the primer occurring for only 1 to 20 nucleotides before dissociation of the enzyme. The order of template preference with respect to processivity was found to be poly(rA), single-stranded heteropolymeric RNA, single-stranded heteropolymeric DNA, double-stranded DNA and poly(dA). In general HIV-1 RT appears to be processive on RNA templates and more distributive in nature on DNA templates.

The high mutation rate observed for the HIV-1 genome, and hence the rapid rate of genome evolution observed for HIV-1, can be attributed to two features of the RT. Firstly, unlike most cellular DNA polymerases, retroviral RTs lack a 3'-5' proofreading capability and therefore any errors made during template copying are not rectified. Secondly RTs, and in particular HIV-1 RT, are notoriously error prone. Studies using several systems have shown that the estimated error rate of HIV-1 RT copying a RNA template *in vitro* is approximately 1 misincorporation per 6,900 nucleotides (Ji and Loeb, 1992). This rate is at least ten times greater than that of any other retroviral RT so far examined. In comparison MuLV RT has an error

rate on a RNA template of approximately 1 in 28,000 and that of *E.coli* DNA polymerase I is approximately 1 in 37,000. The frequency of misincorporation in HIV-1 is amplified by the fact that errors also occur at a rate of approximately 1 in 5900 when the DNA strand is used as a template (Ji and Loeb, 1992; Yu and Goodman, 1992). Another major determinant in the infidelity of HIV-1 RT is the fact that the enzyme has been shown to be capable of extending the DNA residues at mismatched 3'-termini with high efficiency (Peliska and Benkovic, 1992). These features of RT fidelity are sufficient on their own to account for the observed mutation rate of the HIV-1 genome *in vivo*.

1.4.2.2) The RNase H Activity of HIV-1 RT.

RNase H activity was initially detected in extracts of calf thymus during attempts to purify a DNA-dependent RNA polymerase after a factor was found that specifically degraded the RNA moiety of the RNA.DNA hybrid formed by the RNA polymerase (Stein and Hausen, 1969). Since then a large number of RNases H have been described in many different organisms ranging from bacteria to large multicellular animals (Wintersberger, 1990). The biological role of RNase H within the eukaryotic cell is unclear at present although it is possible that the enzyme plays a role in DNA replication (Wintersberger, 1990).

Retroviral RNase H was first identified by Moelling *et al.* (1971) where it was associated with the RT of AMV. The RNase H activity plays a vital role in retroviral reverse transcription.

Schatz *et al.* (1990A) and Tisdale *et al.* (1991) showed that inactivation of the RNase H domain of HIV-1 RT resulted in a complete loss of viral infectivity. This indicated that not only is the retroviral RNase H essential for replication but also that its activity cannot be replaced by cellular counterparts. This observation makes the retroviral RNase H activity a potentially important target for antiretroviral drugs.

As stated previously, the role of RNase H is to digest the RNA moiety of a DNA.RNA hybrid. In reverse transcription this hydrolysis serves two main functions. The first is the removal of the viral RNA genome and the tRNA primer thereby leaving the minus-strand ((-)-strand) DNA free to act as a template for plus-strand ((+)-strand) DNA synthesis and the second is the formation of the primer for (+)-strand DNA synthesis. RNase H activity has been shown to be dependent on the presence of Mg^{2+} with optimum activity occurring at a pH of between 8 - 8.5 and at a salt concentration of 50 mM (Starnes and Cheng, 1989). A distinguishing feature of RNase H mediated cleavage of RNA is that hydrolysis of the phosphodiester bond results in the production of oligoribonucleotides with free 3'-OH and 5'-phosphate groups (Crouch and Dirksen, 1985). This is in contrast to the action of most other nucleases that leave free 5'-OH and 3'-phosphates.

RNase H activity was originally thought to catalyse the hydrolysis of the phosphodiester bond of the RNA moiety as an obligatory 3'-5' exonuclease (Keller and Crouch, 1972; Leis *et*

al., 1973) based on an observation that the retroviral RNase H was apparently unable to cleave the RNA moiety of certain circular RNA.DNA hybrids. The observation that the degradation end products of HIV-1 RNase H are primarily mono- di- and tri-nucleotides was also considered to support the proposal that RNase H acted in an exonucleolytic manner (Starnes and Cheng, 1989). On the other hand, some evidence suggested that retroviral RNase H could also act as an endonuclease with the finding that both the tRNA primer and the (+)-strand primer were removed intact by RNase H (Omer and Faras, 1982; Champoux *et al.*, 1984). Clarification of the nucleolytic mechanism by which retroviral RNase H cleaves RNA was supplied by Krug and Berger (1989). They found that both HIV-1 RT associated RNase H and AMV RT associated RNase H were capable of cleaving relaxed, circular, covalently closed plasmids in which 270 residues of one strand were ribonucleotides. They also showed that the AMV enzyme was capable of deadenylating capped globin mRNA with a covalently attached oligo(dT) tail at the 3′ terminus. These results proved that retroviral viral RNase H activity was in fact endonucleolytic. These findings were later supported by results obtained by a number of groups that indicated that retroviral RNase H is an endonuclease (Oyama *et al.*, 1989). Schatz *et al.* (1990B) presented evidence that HIV-1 RNase H displays both endonucleolytic and 3′- 5′ exonucleolytic activities. They found that the addition of HIV-1 RT to a defined RNA.DNA hybrid substrate, radio-labelled at the 3′ end of the RNA, resulted in a spectrum of RNA fragments being produced that extended from the 3′ label. These fragments differed from each other by one

nucleotide, a result indicative of a 3'- 5' endonuclease activity. Labelling of the 5' terminus of the RNA, however, resulted in the production of only two major cleavage products extending from the 5' end of the RNA, a result suggesting that RNase H also possesses a 3'- 5' exonuclease activity after an initial endonuclease cleavage.

The processivity of the endonuclease activity of HIV-1 RT associated RNase H was examined by DeStefano *et al.* (1991A) who found that the enzyme exhibited a partially processive endonuclease activity. This statement was based on the observation that the enzyme was not processive enough to generate the maximum possible amount of cleavage during a single binding event of the enzyme with the hybrid RNA.DNA substrate. They also found that the predominant mechanism of cleavage was endonucleolytic.

The question as to whether or not the site of cleavage by RNase H is sequence dependent and enzyme dependent was addressed by Mizrahi (1989B). Using 5'-mismatched and fully base paired hybrids as substrates and three different RNases H (*E.coli*, MuLV and HIV-1), she demonstrated that the kinetics of appearance of cleavage products and their distribution differed significantly between the enzymes. This indicated a difference in the sequence dependence of the kinetics of RNase H cleavage between different enzymes and suggested the preference of special sites for hydrolysis. The basis of this difference is still unclear at present but it has been proposed that secondary structures within

the substrate may be recognised by the enzymes (Mizrahi, 1989B).

1.4.2.3) The Spatial and Temporal Relationship of the Polymerase and RNase H Activities.

Although it is known that the RNase H activity and the polymerase activity of HIV-1 RT are functionally but not structurally independent (Jacobo-Molina and Arnold, 1991B), the relationship between the activities during polymerisation is still not entirely clear.

Early estimates of the spatial separation between the two active sites of HIV-1 RT, based upon the size of the cleavage products produced by RNase H activity, varied between 7 and 18 nucleotides. More accurate measurements by Furfine and Reardon (1991) indicated that the separation between the two active sites was 15 to 16 nucleotides. This was based on observations that the RNase H activity hydrolysed a phosphodiester bond of RNA at a position between the 15th and 16th ribonucleotide upstream of the 3'-terminus of a DNA primer which defined the location of the polymerase active site. This cleavage site remained a fixed distance behind the 3'-primer terminus as the polymerase extended the primer. The crystallographic distance data of Arnold *et al.* lent support to this observation (Arnold *et al.*, 1992). The more recent findings of Gopalakrishnan *et al.* (1992), based on a single binding event of the enzyme, indicated that the distance between the two active sites is in fact 18 to 19 nucleotides. The authors attribute this discrepancy to the fact that the primary

18/19 base cut is observable only at an early stage of the reaction, and that further rapid secondary degradation occurs to give rise to the final 15-16 nucleotide fragment observed by Furfine and Reardon (1991). The result appears to be consistent with the more accurate structural data of Kohlstaedt *et al.* who found that the separation of the active sites in relation to an A-form RNA.DNA hybrid was 20 nucleotides (Kohlstaedt *et al.*, 1992).

Two main mechanisms have been proposed regarding the temporal relationship between the polymerase and the RNase H of HIV-1 RT. The first is that the two activities may act in a coordinated fashion during DNA synthesis, i.e. a coupled mechanism. The second is that the RNA is cleaved infrequently, or not at all, during DNA synthesis and that the same or another RT must return at a later stage to perform the bulk of RNA degradation, i.e. an uncoupled mechanism.

As discussed earlier, DeStefano *et al.* found that HIV-1 RNase H made a limited number cleavages during polymerisation which was indicative of a partially processive 3'-5' endonuclease activity (DeStefano *et al.*, 1991A). This means in effect that although polymerisation and RNase H activity appear to be coupled the maximum number of possible cuts is not made by the RNase H. In a second paper by these authors (DeStefano *et al.*, 1991B) it was shown that a number of retroviral RTs cut the RNA at least once during DNA synthesis but that much of the potentially degradable hybrid was still left intact. They concluded,

therefore, that the polymerase and RNase H activities were not strictly coupled during reverse transcription. In contrast, a number of groups have provided evidence that suggests that the polymerase and RNase H activities of the enzyme are coupled. Schatz *et al.* (1990B), whilst investigating the mechanism of RNA degradation by RNase H, observed that cleavage by RNase H occurred concomitantly with the elongation of the DNA oligonucleotide primer. Furfine and Reardon (1991) also made similar observations, finding a temporal correlation between the extension of a DNA oligonucleotide primer by the polymerase activity and the degradation of the RNA moiety by RNase H. The more recent findings of Gopalakrishnan *et al.* (1992) support these observations. This group found that the time dependent growth of a DNA primer from 25 to 27 nucleotides catalysed by the polymerase activity of HIV-1 RT correlated exactly with the time dependent production of RNase H cleavage products of 33, 32 and 31 nucleotides. This information, in conjunction with the structural knowledge that the RNA.DNA duplex lies in groove spanning the polymerase and RNase H active sites, strongly supports the proposal of coupled polymerase and RNase H activities during reverse transcription.

1.4.3) The Process of Reverse Transcription.

The process of reverse transcription is a complex one and all of the steps required are carried out solely by the viral RT. The most generally accepted model for reverse transcription was first proposed by Gilboa *et al.* (1979A) and was based

predominantly on research carried out on the MuLV RT. An important function of reverse transcription is the regeneration of the LTR's of the viral genome. After transcription of the integrated proviral DNA by RNA polymerase II, parts of the terminal sequences, including the 5' promoter region, are absent from the viral RNA genome (Hsu *et al.*, 1978; Shank *et al.*, 1978; Gilboa *et al.*, 1979B; Hu and Temin, 1990). The process of reverse transcription regenerates these terminal regions in the proviral DNA (see Figure 6).

1.4.3.1) Initiation of Reverse Transcription.

Reverse transcription is initiated by the hybridisation of the 3' stem of a specific tRNA molecule to a region near the 5' end of the viral RNA genome known as the primer binding site (pbs). The use of a specific tRNA as a primer for the initiation of polymerisation is common to many retroviral RTs. For instance the MuLV RT uses $tRNA^{Pro}$ for initiation (Harada *et al.*, 1975) whereas the Rous sarcoma virus enzyme employs $tRNA^{Trp}$ (Baltimore, 1970). Sequence analysis of the pbs of HIV-1 indicated that the initiation primer used by HIV-1 RT was human $tRNA^{Lys3}$ (Ratner *et al.*, 1985; Sanchez-Pescador *et al.*, 1985; Wain-Hobson *et al.*, 1985). The specificity of HIV-1 RT for $tRNA^{Lys3}$ was demonstrated by the fact that the heterodimer preferentially bound $tRNA^{Lys3}$ even in the presence of a 100-fold excess of other species of tRNA (Barat *et al.*, 1989). Cross-linking studies have shown that HIV-1 RT specifically binds a synthetic $tRNA^{Lys3}$ via a 12 nucleotide portion of the 3' end of the tRNA, an area

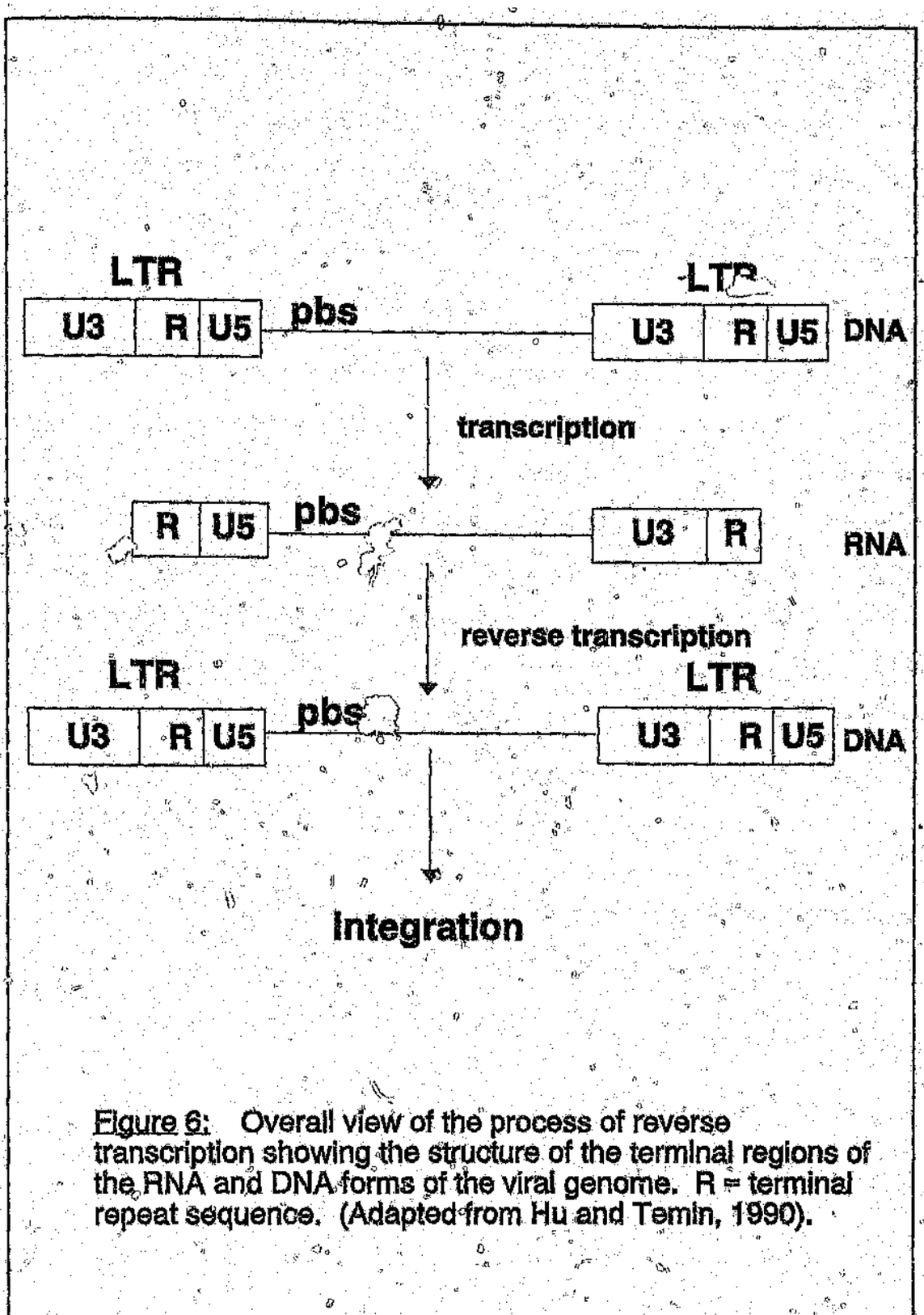


Figure 6: Overall view of the process of reverse transcription showing the structure of the terminal regions of the RNA and DNA forms of the viral genome. R = terminal repeat sequence. (Adapted from Hu and Temin, 1990).

encompassing the anticodon domain (Barat *et al.*, 1989). In addition Sallafranque-Andreola *et al.* (1989) have shown that polymerisation on synthetic substrates is inhibited by the presence of $tRNA^{Lys3}$. As stated earlier both the p51 and p66 subunits have been implicated in the binding of $tRNA^{Lys3}$ (Barat *et al.*, 1989).

Hybridisation of the $tRNA^{Lys3}$ primer has been shown to be aided by the presence of the 7-kD nucleocapsid protein p7(NC) (Prats *et al.*, 1988; Barat *et al.*, 1989). The protein contains a "zinc finger" like motif and has been implicated in a number of interactions with the viral genomic RNA (South *et al.*, 1989). Of particular relevance is the ability of p7(NC) to unwind tRNA (Khan and Giedroc, 1992) and it has been shown that this facilitates the formation of RNA.RNA duplexes as well as increasing the yield of DNA products produced after polymerisation by RT (Prats *et al.*, 1988).

The 18 nucleotide pbs of HTV-1 RT may also play a role in the assembly of viral particles. Nagashunmugam *et al.* (1992) constructed a proviral DNA containing a mutant pbs which, after transfection into a CD4+ cell line, affected viral production giving rise to non-infectious viral particles and abnormal amounts of internal viral structural proteins, in particular p24. Reverse transcription was also abolished. These results suggest a role for the pbs in viral replication that is independent from its function in the initiation of reverse transcription.

1.4.3.2) The Synthesis of (-)-Strand DNA.

Following annealing of the $tRNA^{Lys3}$ primer to the pbs of the RNA genome within the template-primer binding groove of HIV-1 RT, synthesis is initiated at the 3′ end of the $tRNA^{Lys3}$ primer. The pbs of HIV-1 is located at a position close to the 5′ end of the viral RNA, and DNA synthesis proceeds from the primer towards the 5′ end of the RNA with the start of synthesis corresponding very closely to the 3′-terminus of th U5 sequence (Goff, 1990; Gilboa *et al.*, 1979A; Peliska and Benkovic, 1992). The end result of this process is the synthesis of short stretch of (-)-strand DNA complementary to the U5 and direct repeat sequences of the 5′-terminus of the viral genome. This is known as (-)-strand strong-stop DNA (see Figure 7 (step 1)). The RNA template corresponding to this DNA is degraded by the RNase H activity. In order for synthesis of the (-)-strand DNA to be completed, the (-)-strand strong-stop DNA must then be translocated to a site near the opposite 3′ end of the same or another viral RNA molecule, a process guided by the complementary direct repeat sequence found at the ends of both the acceptor RNA molecule and the newly formed (-)-strand strong-stop DNA (Figure 7 (step 2). This process is known as strand transfer (Peliska and Benkovic, 1992). A more detailed description of the possible mechanism of strand transfer will be discussed in section 1.4.3.4.

Transfer of the (-)-strand strong-stop DNA to the 3′ end of the RNA genome enables synthesis of the (-)-strand to be continued with the coordinated degradation of the RNA template

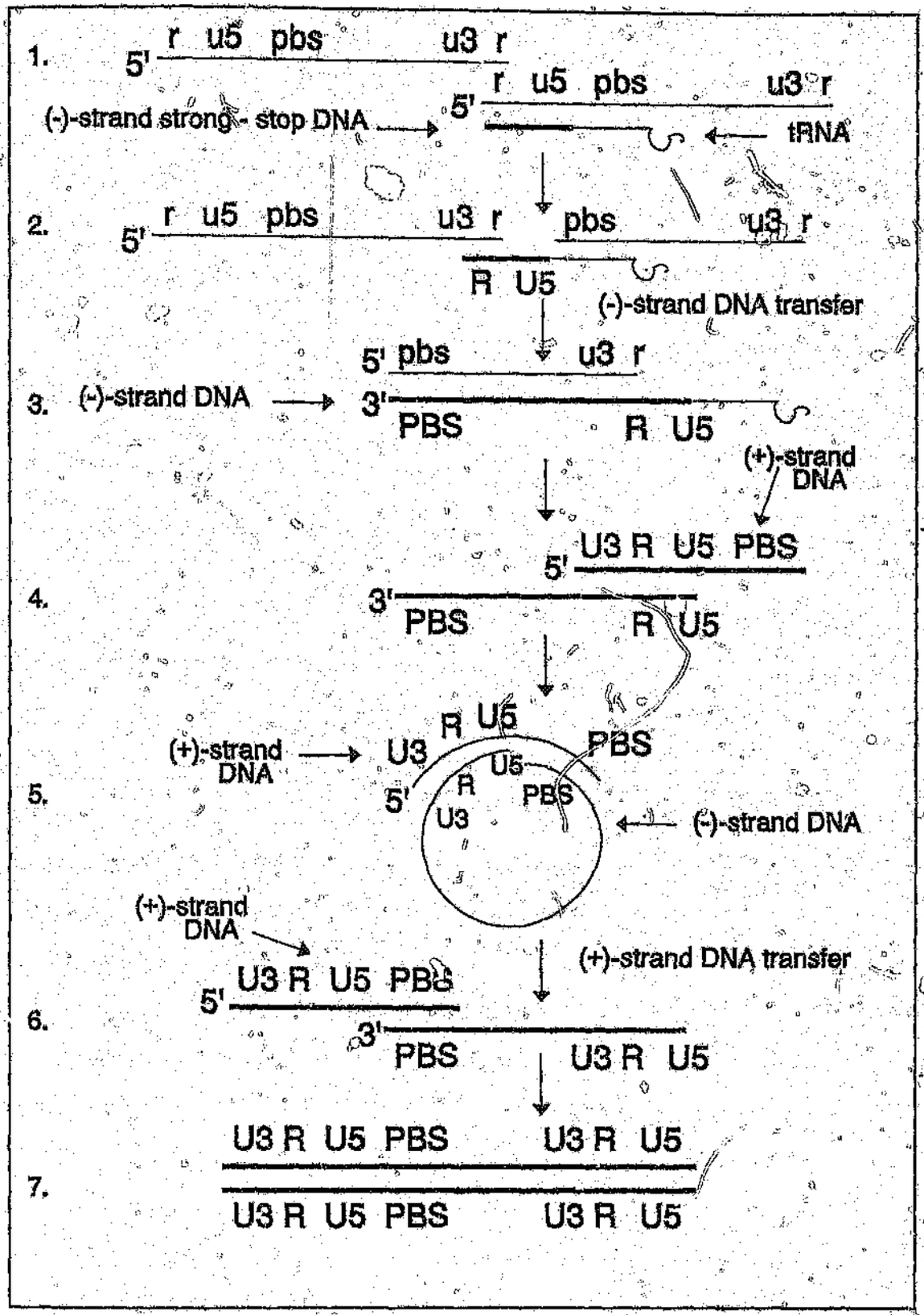


Figure 7: Reverse transcription of a retroviral genomic RNA into a full length proviral DNA by HIV-1 RT. Bold lines denote DNA, thin lines represent RNA. See text for detailed description. (Adapted from Peliska and Benkovic, 1992).

by RNase H (Figure 7 (step 3). However the synthesis of the (-)-strand does not continue until it reaches the 5′ end of the RNA template, but instead appears to stop after the pbs of the RNA template has been copied (Omer *et al.*, 1984).

1.4.3.3) The Synthesis of (+)-Strand DNA.

The synthesis of the second (+)-strand of DNA requires the formation of a specific RNA primer by the RNase H activity from which polymerisation can be initiated. The RNA primer is formed by the selective action of RNase H on the RNA template at a site near the U3 region at the 5′-terminus of the (-)-strand DNA known as the polypurine tract (Taylor and Shameen, 1987; Smith *et al.*, 1984). The polypurine tract appears to be resistant to RNase H degradation. Mutational studies of the polypurine tract region suggest that RT RNase H can recognise a portion of the polypurine tract and then reach over a specific number of bases to make a nick between the U3 region and polypurine tract (Rattray and Champoux, 1987; Rattray and Champoux, 1989). The action of RT DNA-dependent DNA polymerase continues from the 3′ end of the polypurine tract, thereby copying the U3, direct repeat (R), U5 and the $tRNA^{Lys3}$ primer sequences of the (-)-strand template (Figure 7 (step 4)). The copying of the $tRNA^{Lys3}$ primer (which is still attached to the 5′ end of the template) results in the regeneration of the 18 nucleotides of the (+)-sequence for the pbs. The $tRNA^{Lys3}$ primer is then removed by the RNase H activity. This short stretch of (+)-strand DNA, formed from the polypurine tract to the 3′ end of the viral genome, is called the (+)-strand

strong-stop DNA.

As for the (-)-strand, the completion of synthesis of the (+)-strand requires a second template switching event whereby the (+)-strand strong-stop DNA is transferred to the 3' end of the (-)-strand template (Figure 7 (step 5)). Annealing of the (+)-strand strong-stop DNA to the 3' end of the (-)-strand DNA is brought about as a result of the complementarity of the pbs sequences at their 5' and 3' ends respectively. The (+)-strand strong-stop DNA is then extended along the length of the (-)-strand template to complete (+)-strand synthesis. The final step in the formation of the double-stranded proviral DNA is the completion of synthesis of the (-)-strand (Figure 7 (step 6)).

The end result of reverse transcription is the formation of a complete linear, double-stranded, proviral DNA that contains the regenerated LTRs and promoter regions of the virus. This linear proviral DNA has been found to be complete and blunt-ended *in vivo* (Roth *et al.*, 1989; Fujiwara and Craigie, 1989) and, although it is thought that this is the last stage in retrovirus replication in which the RT is required, the proviral DNA must undergo one further modification before integration into the host genome. This last action requires the trimming of two nucleotides from the blunt 3' ends of the (+)- and (-)-strands of the DNA and is mediated by the viral integrase protein.

1.4.3.4) The Mechanism of DNA Strand Transfer During Reverse Transcription.

Since HIV-1 virions each contain two copies of the viral RNA genome, DNA transfers during reverse transcription could theoretically occur either inter- or intra-molecularly. Observations made by Panganiban and Fiore (1988) suggest that the events associated with the transfer of the (-)-strand strong-stop DNA predominantly occur inter-molecularly whereas those associated with the transfer of (+)-strand strong-stop DNA are intra-molecular. Results obtained by Hu and Temin (1990) confirm the nature of the (+)-strand strong-stop DNA transfer but suggest that the (-)-strand strong-stop DNA transfer can occur both inter- and intra-molecularly.

The finding that (+)-strand strong-stop DNA transfer is intra-molecular suggests the formation of a circular structure (see Figure 7 (step 5)). Such circular intermediates have been found after lysis of the virion using detergents (Goff, 1990). This in turn suggests that RT is able to displace the one DNA strand whilst completing synthesis of the (-)-strand.

The fact that (-)-strand strong-stop DNA transfer is able to occur inter-molecularly implies that the RT is able to accommodate both the acceptor RNA molecule and the newly formed RNA.DNA hybrid. Peliska and Benkovic (1992), using a 40/20 base RNA.DNA substrate based on the HIV-1 terminal repeat and a 41 base RNA acceptor molecule (of which 20 bases at the 3' terminus

were identical to the 5' terminus of the primary RNA), found that HIV-1 RT was able to accommodate a RNA.DNA.RNA intermediate. However, transfer of the DNA to the acceptor RNA molecule was found to occur in three stages. Firstly, binding of the initial RNA.DNA substrate to the RT precluded the binding of the RNA acceptor molecule until the DNA had been extended to its full length and a majority of the RNA moiety removed forming a 14/40 base pair RNA.DNA intermediate. The RNA acceptor was then able to bind to the RT forming an RNA.DNA.RNA complex. Secondly, hybridisation of the DNA to the acceptor RNA molecule first required the removal of the small 14 base RNA moiety from the (+)-strand strong-stop DNA by a kinetically distinct action of the RNase H that was not coupled with polymerisation. This demonstrated the necessity of RNase H activity for efficient strand transfer. Finally, once the small RNA fragment had been removed, strand transfer of the acceptor RNA to the (-)-strand strong-stop DNA could occur after which DNA synthesis using the acceptor RNA as a template proceeded processively.

The findings of Peliska and Benkovic (1992) also indicated that RT was capable of extending the DNA by an additional single nucleotide past the 5' end of the primary RNA template. This was found to be independent of the sequence of both the primary and acceptor RNA molecules. This additional nucleotide could be incorporated into the final DNA product and may indicate a further source of error during reverse transcription.

1.4.4) Mutagenesis Studies of HIV-1 RT.

A considerable amount of the information gained regarding the structure and function of HIV-1 RT has been obtained by the use of techniques that introduce defined mutations at specific conserved sites within the enzyme. The types of mutations employed in these studies have varied from large insertions and deletions (Hizi *et al.*, 1988; Prasad and Goff, 1989) to single point mutations of individual amino acids (Mizrahi *et al.*, 1990; Schatz *et al.*, 1989; Larder *et al.*, 1987). Comparative protein sequence analysis of a large variety of retroviral polymerase and RNase H sequences, as well as related sequences from other sources, has been carried out by a number of groups (Johnson *et al.*, 1986; Doolittle *et al.*, 1989). This resulted in the identification of a number of highly conserved amino acid residues and motifs. Jacobo-Molina and Arnold (1991) organised these conserved residues into thirteen different regions of which eight were found within the polymerase domain (regions 1 to 8) and five within the RNase H domain of the protein (regions 9 to 13) (see Table 1). The conserved regions presumably corresponded to areas that played important structural and/or functional roles in the enzyme. The authors designated a region as being conserved according to three criteria. Firstly, that at least half of the amino acid positions within a region, when aligned, had to share identity with the consensus residue at that position (the consensus sequence was based on that of the HIV-1 III_B isolate); secondly, that insertions or deletions occurred only rarely within a region for any sequence examined; and finally, regions

Region	Residue Number	Sequence
1	23 to 37	QWPLTEEKIKALVEL
2	55 to 81	PYNTPVFAIKKDSTKWRKLVDFRELN
3	102 to 134	KKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPS
4	143 to 170	RYQYNVLPQGWKGSPAIFQSSMTKILEP
5	178 to 191	IVIYQYMDDLYVGS
6	211 to 222	RWGLTTPDKKHQ
7	227 to 237	FLWMGYLHPD
8	257 to 256	IQKLVGKLNW
9	441 to 445	YVDGA
10	475 to 486	QKTELQAIYLAL
11	493 to 501	VNIVTDSQY
12	531 to 542	VYLAWVPAHKGI
13	544 to 552	GNEQVDKL

Table 1: Conserved regions within the polymerase and RNase H domains of retroviral RTs. Residues that are conserved in 80% or more of the aligned retroviral sequences are underlined. Residues are numbered according to the HIV-1 RT sequence. (Adapted from Jacobo-Molina and Arnold, 1991)

that were previously defined by other groups were retained so as to provide a degree of consistency and continuity.

The majority of early mutagenesis studies of HIV-1 RT provided information relating to the structural and function l relationship of the two catalytic domains of the enzyme. As discussed previously, it was found that although the two domains appear to be structurally independent, there is a large degree of functional interdependence between the two activities since certain mutations within one domain were able to seriously affect the activity of the other domain (Schatz *et al.*, 1989; Mizrahi *et al.*, 1990). These mutagenesis studies also provided information relating to the specific role of a conserved region or amino acid residue in the structure or function of the enzyme. In summarising these data Jacobo-Molina and Arnold (1991) were able to assign potential functions to the various conserved regions that they had defined based upon the sequence data. Many of these findings were subsequently confirmed with the publication of the crystal structure of the HIV-1 RT heterodimer (Kohlstaedt *et al.*, 1992) and the HIV-1 RNase H domain (Davies *et al.*, 1991), especially with respect to the structural role of certain conserved residues. This was illustrated particularly well by the elucidation of the role played by the universally conserved histidine residue at position 539 (H539). As discussed previously, mutagenesis studies had previously indicated that this residue was very important for full RNase H activity in the enzyme (Schatz *et al.*, 1989; Tisdale *et al.*, 1990), with mutations reducing both substrate binding efficiency and catalytic rate;

however, the exact nature of its role was unclear. With the publication of the crystal structures it was found that this residue plays an important role in stabilising the structure of the RNase H domain through an interaction with the p51 subunit of the heterodimer thereby also forming a portion of the substrate binding site. However, although the knowledge of the detailed structure of HIV-1 RT has largely superseded the use of mutagenesis as a means of investigating the structure of the enzyme, mutagenic techniques are still applicable to the investigation of the catalytic activities associated with reverse transcription.

1.4.4.1) Mutagenesis Studies of the Active Site of the Polymerase Domain of HIV-1 RT.

The conserved motif in HIV-1 RT which is thought to form the active site of the polymerase domain has the sequence Tyr183-Met184-Asp185-Asp186 (YMDD) and is found in region 5 as defined by Jacobo-Molina and Arnold (1991) (see Table 1). The YMDD sequence is highly conserved among many RNA-dependent DNA polymerases although the second residue of the motif does show some variation among other RTs and RT-containing elements (Argos, 1988; Kamer and Argos, 1984). The RT polymerase motif does show some similarity to the DNA-dependent DNA polymerase signature sequence, YGDTDS, suggesting a common catalytic mechanism during polymerisation (Wong *et al.*, 1988). A number of groups have investigated this motif in HIV-1 RT using mutagenesis techniques, and have found that the conservative mutation of either Asp185

or Asp186 completely abolishes the polymerase activity of the enzyme while not affecting the RNase H activity (Larder *et al.*, 1989A; Larder *et al.*, 1987). Mutation of the tyrosine residue at position 183 to a serine resulted in a marked reduction in the enzyme's polymerase activity (Larder *et al.*, 1987). In a recent study Wakefield *et al.* (1992) examined the role of Met184, the least conserved of the four residues of the motif, and found that this residue could be substituted with either a valine, alanine or serine residue without deleterious effects to the enzyme function. This finding was consistent with sequence data from other RTs which has shown that these residues are found at this position. For instance the RTs from MuLV and FLV contain YVDD (Donahue *et al.*, 1988; Shinnick *et al.*, 1981) whereas RTs associated with *E.coli* and *M.xanthus* have the sequence YADD at this position (Lampson *et al.*, 1989; Inouye *et al.*, 1989). However, other mutations at Met184, for instance the insertion of tyrosine (Larder *et al.*, 1989A), glycine and proline (Wakefield *et al.*, 1992), were found to drastically reduce the polymerase activity. In general the results obtained in mutagenesis studies suggest that the YMDD motif plays a crucial catalytic role in the polymerase activity of HIV-1 RT with Met184, the second conserved residue, possibly being involved in structural interactions with neighbouring amino acids. On the basis of molecular modelling studies this motif has been postulated to be at or very near the polymerase active site and may be involved in template recognition and/or divalent metal ion binding.

Although the YMDD motif is the most distinct of the conserved sequences within the polymerase domain and appears to form much of the active site, another conserved residue has been implicated in the catalysis of polymerisation, namely the conserved aspartic acid residue found at position 110 (Asp110) within region 3 (see Table 1). This residue is also conserved among all polymerases and is structurally analogous to the aspartic acid found at position 705 in the Klenow fragment of *E.coli* DNA polymerase 1. The conservative mutation of this residue to glutamic acid (Larder and Kemp, 1989B) also results in a severe depletion or loss of polymerase activity whilst not affecting the RNase H activity. The similarities between the gross structure of the HIV-1 RT polymerase domain and that of the Klenow fragment of *E.coli* DNA polymerase I, and fact that the two enzymes share some sequence homology at the putative polymerase active site (Asp110, Asp185 and Asp186 in HIV-1 RT, Asp705, Asp882 and Asp883 in the Klenow fragment), suggests that HIV-1 RT employs the same catalytic mechanism as *E.coli* DNA polymerase 1 during polymerisation.

1.4.4.2) Mutagenesis Studies of the Active Site of the RNase H Domain of HIV-1 RT

In the case of the RNase H activity of HIV-1 RT, four residues, that are conserved among all known RNase H enzymes, are thought to be associated with the active site. These are the acidic residues Asp443, Glu478 and Asp498 and Asp549. As discussed previously, the crystal structures of both the HIV-1

RT RNase H domain (Davies *et al.*, 1991) and *E.coli* RNase HI (Katayanagi *et al.*, 1990;1992) show these four residues to be in close proximity to each other and capable of binding either one or two divalent metal ions. Kanaya *et al.* (1990), in an initial study of the catalytically important amino acids of RNase H, carried out site-directed mutagenesis of conserved residues in *E.coli* RNase HI (Asp10, Glu48 and Asp70) that correspond to the residues Asp443, Glu478 and Asp498 of the HIV-1 RT RNase H domain. They found that the individual mutation of the conserved aspartic acid residues to either asparagine or glutamic acid resulted in a marked decrease or loss of the RNase H activity of the mutant enzyme. Likewise the mutation of Glu48 to glutamine and aspartic acid also resulted in the loss of the RNase H activity. These findings identified these three acidic residues as being vital for catalysis *E.coli* RNase HI.

Since the HIV-1 RNase H domain shows large similarities to the *E.coli* RNase HI enzyme at a structural level (although at a sequence level the overall homology is low) (see Figure 4), it is reasonable to expect that the seven conserved residues that are shared between the two proteins play similar roles in catalysis by RNase H. Schatz *et al.* (1989) used site-directed mutagenesis techniques to investigate the Glu478 residue of HIV-1. By substituting the glutamic acid at this position for a non-ionisable glutamine residue this group were able to create a mutant enzyme that exhibited a defective RNase H activity whilst maintaining a normal polymerase function.

In a series of publications, Mizrahi *et al.* have investigated, in detail, the role of the two conserved aspartic acid residues, Asp443 and Asp498, in the RNase H activity of HIV-1 RT. Using site-directed mutagenesis, Mizrahi *et al.* (1990) created three individual mutants within the RNase H domain. The first mutation was of the Asp443 residue to asparagine (D443N). This was a conservative mutation whereby an acidic, ionisable carboxylate residue was effectively changed to a non-ionisable, isosteric carboxyamide group. This mutation was therefore expected to cause no structural alteration in the enzyme. Upon expression and characterisation of this D443N mutant it was found that the enzyme retained normal wild type polymerase activity but lacked any RNase H activity. This finding was consistent with those of Kanaya *et al.* (1990) who carried out the equivalent mutation (D10N) on the *E. coli* enzyme with similar results. This result suggested a vital role for the carboxylate group of Asp443 in RNase H catalysis. The analogous mutation of Asp498 (D498N) was then generated. This mutation gave rise to an unstable enzyme that could not be purified or characterised, a result suggestive of an important structural as well as a catalytic role for Asp498. The third mutant enzyme created by Mizrahi *et al.* was a double mutant consisting of the substitution of both aspartic acids by asparagine (D443N/D498N). Surprisingly the distal D443N mutation was found to stabilise the D498N mutation giving rise to a structurally stable mutant enzyme that lacked RNase H activity but retained a wild-type polymerase activity. Two models have been proposed to explain this finding in the double mutant. Firstly it was suggested that in the case of the single D443N

mutant, differences between the hydrogen bonding properties of the carboxylate group of aspartic acid and the carboxyamide of asparagine may cause slight alterations within the local hydrogen bonding or electrostatic network of the protein. These changes however are not severe enough to cause any gross disruption of the protein structure but may, in the case of the double mutant, be sufficient to accommodate the folding defects observed in the single D498N mutant. Secondly the increased stability of the double mutant relative to the D498N mutant is probably the result of an artificial hydrogen bonding interaction between the two juxtaposed asparagine residues in the double mutant and therefore not necessarily a true reflection of any structural relationship, such as a Mg^{2+} bridge between Asp443 and Asp498 in the wild-type enzyme.

Further investigation of the two RNase H$^-$ mutants created in the first study was carried out by Dudding *et al.* (1991). Initial investigations of the polymerase activity of the mutant enzymes suggested that they possessed levels of activity similar to that of the wild-type enzyme (Mizrahi *et al.*, 1990). However a more detailed examination of the RNA- and DNA-dependent DNA polymerase activities of the two mutants on homopolymer and heteropolymer substrates revealed some differences between the wild-type and RNase H$^-$ enzymes. Dudding *et al.* (1991) found that the steady-state kinetics and processivity of the wild-type and mutant enzymes were the same with respect to their ability to copy a homopolymer poly(rA).oligo(dT) template-primer. In addition, when a DNA heteropolymer template derived from the *gag*

region of HIV-1 was used as a substrate, both the pause sites and the product distribution of DNA synthesis were found to be essentially identical. However, when an RNA template of the same nucleotide sequence as the DNA template was used as a substrate for the enzymes, it was observed that although the mutant and wild-type enzymes shared the same pause sites during DNA synthesis, the respective product distributions were significantly different. This result indicated that the RNase H⁻ mutants were unable to efficiently extend prematurely terminated nascent primer DNA chains, and suggested that the RNase H activity of HIV-1 RT was important for the re-initiation of DNA synthesis on an RNA template after the enzyme had stopped synthesis at sequence determined pause sites. A possible explanation for this observation is that the continuity of efficient extension of DNA on a RNA template by RT is dependent on the cleavage of the RNA template by the RNase H activity at specific sites. This in turn could indicate that the enzyme can be acting either as a polymerase or as a RNase H, but not both simultaneously. However this would not necessarily mean that the enzyme must first dissociate from the RNA template and rebind before RNA cleavage can occur. It is possible that the enzyme may have to undergo a conformational change whilst still attached to the RNA template in order to enable the RNase H to cleave the RNA. After cleavage the enzyme would then have to revert to the polymerase conformation before DNA synthesis could continue. Since they are unable to cleave the RNA template, the RNase H⁻ mutants may become stuck in the RNase H conformation thereby inhibiting further extension of the nascent DNA primer and

requiring the enzyme to dissociate from the RNA template and rebind before any further DNA synthesis can occur.

1.5) AIMS.

The aim of this project is to use the technique of site-directed mutagenesis to create further mutants at the conserved Asp443 and Asp498 positions that would provide a greater insight into the catalytic and structural role of these residues in the RNase H activity of HIV-1 RT, and lead to the proposal of a possible catalytic mechanism for the cleavage of the RNA moiety of a DNA.RNA hybrid by RNase H.

As discussed in the previous section, the work of Mizrahi *et al.* (1990) showed that Asp443 and Asp498 were vital to both the RNase H activity and the structural integrity of HIV-1 RT. The findings also suggested that the carboxylate group of these acidic residues was the factor that was important for both of these features. However, the finding that the distal D443N mutation was able to stabilise the folding defect caused by the D498N mutation suggested the possibility of an important structural relationship between the two residues in the wild-type enzyme.

In order to address the questions relating to the nature of the possible interaction between these two residues, and to further clarify the role of the carboxylate groups in catalysis by RNase H, two mutant RT enzymes have been created using site-directed mutagenesis. In the first mutant Asp443 is substituted for an alanine residue (D443A). The side chain of the alanine residue consists of a methyl group which is sterically smaller

than the carboxylate group of aspartic acid. The methyl group is also chemically inert and therefore unable to be involved in interactions with any other residues or metal ions within the region. Expression and characterisation of the D443A mutant should provide information relating to: 1) The role of the carboxylate group of Asp443 in catalysis by RNase H. 2) The size and charge of the side-chains that can be accommodated at the 443 position without severely disrupting the tertiary protein structure.

In the second mutant enzyme Asp443 and Asp498 are substituted for alanine and asparagine respectively (D443A/D498N). The purpose of this double mutant is to determine the nature of the stability of the D443N/D498N mutant, given that the single D498N mutant is structurally unstable. The replacement of asparagine by an alanine residue at position 443 in the double mutant will prevent any possible interaction with the asparagine residue at position 498. Expression and characterisation of this mutant will provide information relating to the structural role of Asp498 in HIV-1 RT as well as providing further information about the role of the carboxylate groups in catalysis by RNase H.

This project forms part of a larger study in which the roles of the conserved residues Asp443, Glu478, Asn494 and Asp498 are being investigated by site-directed mutagenesis (C.Nkabinde and V.Mizrahi, unpublished data). In the longer term it is hoped that the information gained from this study will contribute to

a greater understanding of the mechanism of the RNase H activity of HIV-1 RT and aid in the identification of the important structural features of this essential activity. This in turn may lead to the development of novel therapeutic agents against HIV-1 through the process of rational drug design.

2.) MATERIALS AND METHODS.

2.1.) Materials.

2.1.1) Reagents.

[Methyl-^{3}H] thymidine 5'-triphosphate (41 Ci/mmol), [α-^{35}S] deoxyribose adenosine triphosphate (>800 Ci/mmol), [α-^{32}P] adenosine triphosphate (>3000 Ci/mmol), and [α-^{32}P] uridine triphosphate were obtained from Amersham International, Amersham, UK. Restriction enzymes, T4 DNA ligase, T4 DNA polymerase, T4 gene 32 protein, isopropyl-α-D-galactopyranoside (IPTG), 5-bromo-4-chloro-3-indolyl-α-D-thiogalactopyranoside (X-gal), lysozyme, polyadenylic acid (poly(rA)), deoxyribonucleotide triphosphates (dNTPs), ribonucleotide triphosphates (rNTPs), dithiothreitol (DTT), ribinuclease A (RNase A), chloramphenicol, ampicillin, molecular weight markers II, III, and VI, alkaline phosphatase and leupeptin were obtained from Boehringer-Mannheim, Mannheim, Germany. T7 RNA polymerase, *E.coli* RNase HI and ribonuclease inhibitor (RNasin) were obtained from Promega, Madison, Wisconsin, USA. Nalidixic acid, uridine, pepstatin, ethidium bromide, amberlite resin beads, ethylenediamine tetra-acetic acid (EDTA), ammonium sulphate, triton X-100, boric acid, trizma base (Tris), urea, phenylmethylsulphonyl flouride (PMSF) and thiamine were obtained from Sigma, St.Louis, Missouri, USA. N,N'-methylene-bis-acrylamide (bis), N,N,N',N'-tetramethyl ethylenediamine (TEMED), acrylamide, ammonium persulphate, agarose, and xylene cyanol were obtained from Bio-Rad, Richmond,

California, USA. dT_{12-18} was obtained from Pharmacia, Upsala, Sweden. Ammonium acetate, calcium chloride, magnesium chloride, D(+)-glucose, diethyl ether, polyethylene glycol 2000 and phenol were obtained from Merck, Darmstadt, Germany. Ethanol, hydrochloric acid, magnesium sulphate, di-sodium hydrogen orthophosphate, sodium acetate, 2-mercaptoethanol, methanol, dimethyl formamide, dimethyl-dichloro silane and 1,1,1,trichlorethane were obtained from BDH, Poole, UK. Isopropyl alcohol and sodium chloride were obtained from Lab Chem, Boksburg, South Africa. Acetic acid, sodium chloride and sodium citrate were obtained from SMM Chemicals, Johannesburg, South Africa. Tryptone and yeast extract agar were obtained from Biolab, Johannesburg, South Africa. Bacto-agar was obtained from Difco, Detroit, Michigan, USA. Yeast extract was obtained from Oxoid Ltd, Basingstoke, UK. Chloroform was obtained from Associated Chemicals, Johannesburg, South Africa.

2.1.2) Oligonucleotides.

All of the oligonucleotides were synthesised using a Beckman System 1 plus DNA synthesiser and were purified by HPLC before use. The following oligonucleotides were used: the mutagenic oligonucleotides D443A (5'-TTCTATGTAGCAGGGGCA-3') and D498N (5'-ATAGTAAACAAATTACCAA-3'), the primers SP1 (5'-ATTCCTGAGTGGGAG-3') and SP2 (5'-TTCTATGTAAATGGGGCA-3'), which were used in conjunction with D443A and D498N for DNA sequencing, and the hybrid substrate oligonucleotide for RNase H assays, RP1 (5'-GTCACATAGTCTCTAAAA-3').

2.1.3) Bacterial Strains and Plasmids.

The *E.coli* strain CJ236 (dut⁻, ung⁻), used for oligonucleotide-directed mutagenesis, was kindly provided by Dr T.Kunkel. The *E.coli* strain JM101 was used for M13 cloning procedures. The *E.coli* strain AR120, which was used as the host strain for the expression of the recombinant RT, was obtained from Dr C.Debouck (SmithKline Beecham Research and Development, USA). The *E.coli* strain DH5α was used for the propagation of the transcription vector pGEM-ΔGAG-A. The RT expression plasmid, pGALKRTE, and the protease plasmid, pDFTPRO4, were as previously described (Mizrahi *et al.*, 1989A). The transcription vector pGEM-ΔGAG-A, which was used for the preparation of (+)-GAG^{345} RNA, was previously described (Dudling *et al.*, 1990).

2.2) Methods.

2.2.1) Oligonucleotide-directed Mutagenesis.

The point mutations of the HIV-1 RT coding region in M13-RT, those of aspartic acid to alanine at position 443 (D443A) and the second site mutation of aspartic acid to asparagine at position 498 (D443A/D498N), were introduced according to a modified version of the method of Kunkel *et al.* (1987). Briefly, the modification was that M13-RT containing the wild-type RT sequence (M13-RT^{wt}), or in the case of the double mutant the D443A background mutation (M13-RT^{D443A}), was passaged twice through CJ236 (dut⁻,ung⁻). This particular *E.coli* strain lacks the

enzymes dUTPase (dut^-) and uracil N-glycosylase (ung^-). dUTPase is responsible for the catabolism of dUTP and lack of the enzyme results in elevated levels of dUTP which competes with dTTP for incorporation into the DNA template during replication. Uracil N-glycosylase is responsible for the removal of uracil that has been incorporated into DNA. Therefore, in the combined dut^-,ung^-, dUTP is incorporated into DNA in place of dTTP which results in a uracil-rich M13-RT DNA template. It was found that the double passaging of M13-RT through CJ236(dut^-,ung^-) increased the uracil content of the M13-RT template DNA which in turn led to an increase in the mutagenesis frequency. The degree of uracil incorporation was assessed by comparing the titre of uracil-rich M13-RT in the permissive host, CJ236(dut^-,ung^-), versus a non-permissive, dut^+,ung^+ strain, JM101. The titre ratio was typically found to be 10^5-10^6. Titres below this value gave mutagenesis frequencies that were too low for rapid identification and isolation of the mutant.

Having obtained sufficient uracil incorporation into M13-$RT^{wt/D443A}$, single-stranded, uracil-containing DNA was isolated from 20 ml of phage-containing supernatant according to a modified version of the method of Sambrook *et al.* (1989A). In order to obtain high quality single-stranded DNA (ss-DNA) for mutagenesis, this procedure was modified to include two rounds of extraction with 1 volume of phenol/chloroform (1:1) followed by extraction with 0.5 volumes of chloroform and 1 volume of diethyl ether. The ss-DNA was then precipitated with 2.5 volumes of ice-cold 100% ethanol and collected by centrifugation at 8000 x g for 30

minutes in a Hägar microfuge (Hägar, Cape Town, South Africa). The DNA was resuspended in TE buffer (pH 8) (10 mM Tris.HCl (pH8), 1 mM EDTA (pH8))to give a final ss-DNA concentration of ± 0.5 µg/µl.

The rest of the mutagenesis procedure was carried out according to the method of Kunkel *et al.* (1987) with the inclusion of one additional step whereby the extent of endogenous priming that occurs during the *in vitro* polymerisation step was assessed (this results in the formation of wild-type M13-RT clones). This was done by setting up a parallel dummy polymerisation reaction that did not contain the mutagenic oligonucleotide primer. The degree of endogenous priming was initially determined by analysis of the polymerisation products on a 1% agarose gel and by the number of plaques obtained after transformation of JM101 with the quenched polymerisation mixtures. Ratios of the number of plaques obtained for the primed and non-primed mixtures were typically 20:1 respectively. The modification of the mutagenesis procedure gave a mutagenesis frequency of approximately 55%.

2.2.2) 5'-[^{32}P]-end Labelling of Oligonucleotides.

Oligonucleotides were 5'-end labelled by incubating 2 pmole of the oligonucleotide with 10 mM DTT, 30 µCi [γ-^{32}P]-ATP (1 pmol), 10 units (according to the manufacturer's definition) of T4 polynucleotide kinase and 1 µl of 10 x polynucleotide kinase buffer in a final volume of 10 µl at 37°C for 45 minutes followed

by heat-inactivation at 65°C for 10 minutes. Incorporation of ^{32}P label into oligonucleotide was monitored by thin layer chromatography (PEI cellulose, developed with 0.3 M potassium phosphate, pH7.0). The percentage incorporation of label varied between 40-80%.

2.2.3) Screening of Mutant M13-RT Clones.

Following oligonucleotide-directed mutagenesis, 24 small, well-spaced plagues (labelled DApl1 to 24 or DA/DNpl1 to 24), were selected for screening using the 5'[^{32}P]-end labelled mutagenic oligonucleotide that was used for the creation of the particular mutant, *i.e* the D443A oligonucleotide was used to screen for the D443A mutation and the D498N oligonucleotide was used for the D443A/D498N mutation. Each plaque was placed in a sterile 15 ml culture tube (Lasec, Johannesburg, South Africa) containing 1.5 ml of Luria broth (LB) (see Appendix B) inoculated with a 1:100 dilution (15 µl) of an overnight culture of JM101. The tubes were then placed in a Series 25 controlled environment shaker/incubator (New Brunswick Scientific Co., Edison, New Jersey, USA) and incubated at 37°C for 5 hours with vigorous aeration (400 rpm). Following incubation, the cultures were placed in 1.5 ml microfuge tubes (Lasec, Johannesburg) and centrifuged at 9000 x g for 10 minutes in order to obtain the phage-containing supernatant. This supernatant was used for the isolation of ss-DNA as described previously. 200 µl of this supernatant was set aside as a phage stock for later use. The ss-DNA was resuspended in 10 µl of TE buffer (pH 8) and 5 µl of this

solution was analysed on a 1% agarose gel (Sambrook *et al.*, 1989B), using a TAE electrophoresis buffer [0.04 M Tris.acetate (pH 8), 1 mM EDTA (pH 8)] and 3 µl of a sucrose based sample loading buffer (0.25% bromophenol blue, 40% (w/v) sucrose). The samples were electrophoresed at 80V for 1-2 hours and the ss-DNA obtained was checked for deletions. If no deletions were evident, then 2.5 µl of each resuspended DNA as well as a negative control (0.5 µg of M13-RT^{wt} ss-DNA) was spotted onto a Biodyne nitrocellulose transfer membrane (Pall Ultrafine Filtration Corporation, Glen Cove, New York, USA) and allowed to air dry. The membrane was prepared by consecutive 2 minute washes in petri dishes containing 0.5 M NaOH/ 1.5 M NaCl and 0.5 M Tris.HCl (pH 8)/1.5 M NaCl. This was followed by rinsing in 2 x SSC (see Appendix B) and air drying. The air dried membrane was then baked at 80°C for 2 hours prior to hybridisation.

Hybridisation was carried out according to standard protocols. The membrane was pre-hybridised by incubation in a solution containing 1.5 ml 20 x SSC, 1 ml 50 x Denhardt's solution (see Appendix B), 2.5 ml H_2O at 37°C for 1 hour. 10 µl of the appropriate labelled mutagenic oligonucleotide was then added and hybridisation was carried out at room temperature for 1 hour. After hybridisation, the probe solution was removed and stored for later use. The membrane was washed in three changes of 6 x SSC for 1, 5 and 5 minutes, sealed in a bag and exposed to 3M XDA Trimax X-ray film (X-ray Imaging Services, Rustenburg, South Africa) for 1 hour at -70°C and developed. Under these initial hybridisation conditions, the oligonucleotide probe

hybridised to both wild-type and mutant versions of the ss-DNA samples present on the membrane. In order to identify those clones carrying the required mutation, the membrane was washed for a further five minutes with 6 x SSC at a temperature approximately 1°C higher than the estimated melting temperature (T_m) of the mutagenic oligonucleotide to its complementary sequence. The T_m was estimated using the formula $T_m = 4(G+C)+2(A+T)$ where G,C,A and T= the number of times guanine, cytosine, adenine and thymine respectively occur in the mutagenic oligonucleotide. The T_m for the D443A oligonucleotide was thus calculated to be approximately 64°C and that for the D498N oligonucleotide was 48°C. Following washing at the appropriate temperature, the membrane was again exposed to X-ray film for 1 hour at -70°C and developed. Those clones still showing hybridisation to the oligonucleotide after washing under these high stringency conditions were identified as possessing the required mutation.

2.2.4) Construction of the Mutagenised HIV-1 RT expression System.

2.2.4.1) Excision of Mutagenised *Asp*718 Fragment.

The presence of the correct mutation in the M13-RT clones identified by screening was confirmed by DNA sequencing using an Amersham M13 dideoxynucleotide sequencing kit (Amersham International, Amersham, UK) according to the manufacturer's protocol. The primers SP1, SP2, and D498N were used / sequence-

the region of M13-RT between nucleotides 1468 and 1835 of the *Bam*H1/*Xba*1 insert, which flank the mutated Asp443 (nucleotides 1522-1524) and Asp498 (nucleotides 1687-1689) codons (see Figure 8). Once the mutagenic sequence had been confirmed, the M13-RT clone selected was grown up in 30 ml of LB containing a 1:100 dilution of JM101 for 5 hours with vigorous aeration (400 rpm). The bacterial pellet was recovered by centrifugation in a 30 ml polypropylene centrifuge tube (Beckman, Palo Alto, California, USA) at 12,000 x g for 10 minutes in a Beckman J2-21 centrifuge using a JA20 rotor. This pellet was then used for the isolation of the double-stranded replicative form DNA (RF-DNA) according to the method of Sambrook *et al.* (1989C) for the rapid isolation of plasmid DNA. The RF-DNA was resuspended in 60 µl of TE buffer (pH 8) and 10 µl aliquots of this DNA was digested with the restriction endonuclease *Asp*718. *Asp*718 digests were typically carried out in a final volume of 12.2 µl consisting of 10 µl of DNA containing solution, 10 units (1 µl) of *Asp*718 and 1.2 µl of 10 x restriction endonuclease buffer B (as supplied by Boehringer-Mannheim) followed by incubation at 37°C for 1 hour. Digestion of M13-RT^{m} RF-DNA with *Asp*718 excised the mutagenized portion of the *Bam*H1/*Xba*1 insert by digestion at nucleotide positions 1476 and 1804 to give rise to a 328-bp fragment (Figure 8). The restriction endonuclease digests were run on a 1% agarose gel at 80 V for 1-2 hours. The bands corresponding to the 328-bp fragment were excised from the gel using a clean scalpel blade and the fragment was recovered from the gel slice by electroelution.

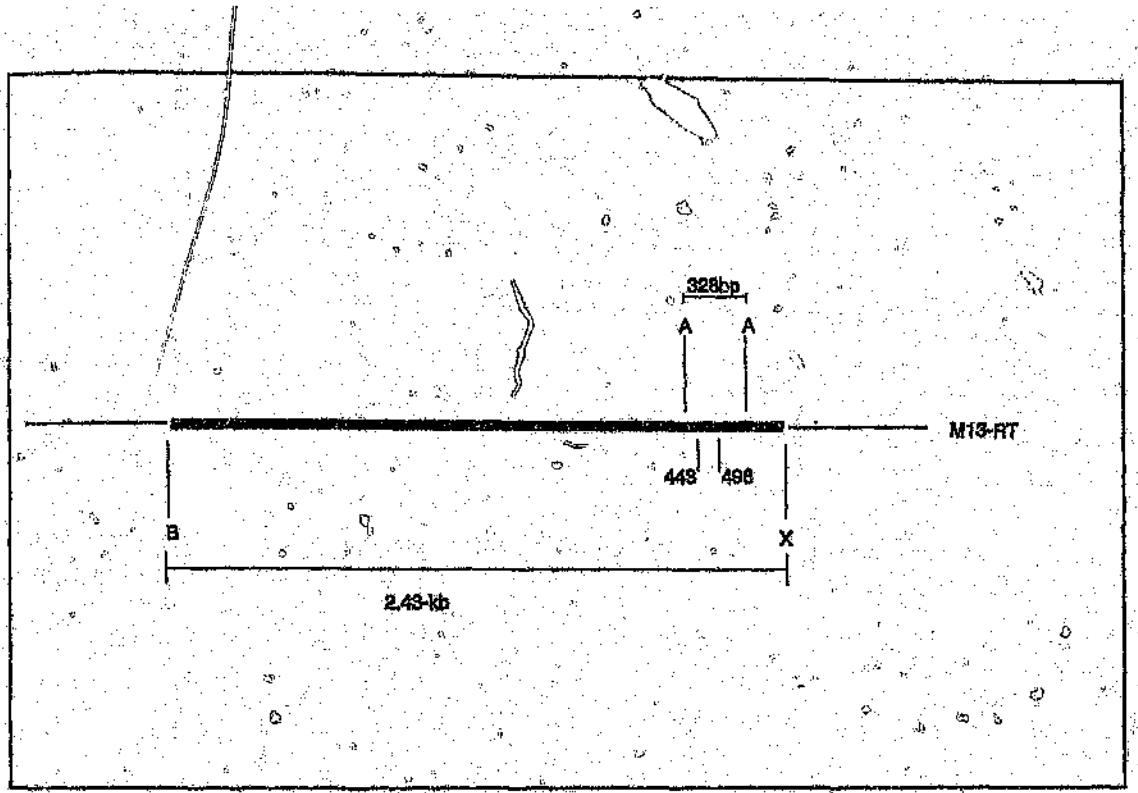


Figure 8: Restriction map of the M13-RT vector. The *Bam*H1/*Xba*1 RT coding fragment is indicated by a bold line. The relative positions of the Asp443 and Asp498 residues are indicated. A= *Asp*718, B= *Bam*HI and X= *Xba*I.

2.2.4.2) Electroelution.

Electroelution was carried out using a Model UAE Unidirectional Electroelutor (International Biotechnologies, New Haven, Connecticut, USA). The procedure involved immersing the gel slice in 20 mM Tris.HCl, 5 mM NaCl and 0.5 mM EDTA and passing a 100 V, 30 mA current through the slice for 30 minutes. Eluted DNA was collected in a well containing a cushion of 100 µl of 7.5 M CH_3COONH_4/bromophenol blue which was then removed along with 300 µl of the surrounding buffer. The bromophenol blue dye was removed by extraction with 1-butanol. The DNA fragment was precipitated by the addition of 1 ml of ice-cold 100% ethanol and collected by centrifugation at 9000 x g for 30 minutes. The DNA pellet obtained was washed with 70% ice-cold ethanol and centrifuged for a further 10 minutes. The DNA pellet was dried and resuspended in 20 µl of TE buffer (pH 8).

2.2.4.3) Reconstruction of M13-RT^m.

The electroeluted, mutagenised 328-bp *Asp*718 fragment was recloned into M13-RT^{wt} from which the original 328-bp wild-type *Asp*718 fragment had been removed (M13-ΔRT). The purpose of this cloning step was to obviate the need to sequence the entire RT coding region for non-specific second-site mutations that may have arisen during the mutagenesis procedure. The vector M13-ΔRT was prepared in a similar manner to that of the *Asp*718 fragment. Briefly M13-RT^{wt} was digested with *Asp*718, and the digest run on a 1% agarose gel. After electrophoresis, the large M13-RT

fragment, lacking the wild-type *Asp*718 fragment, was excised from the gel and isolated by electroelution. The resulting M13-ΔRT vector was then treated with calf intestinal phosphatase so as to prevent self-ligation.

Ligation of M13-ΔRT to the mutagenised *Asp*718 fragment was carried out as described in Appendix A. This ligation mixture was used to transform competent JM101 cells which were prepared according to a modified version of the method of Sambrook *et al.* (1989D). An additional step was added to this method in which the cells were initially gently resuspended in 10 ml of ice-cold 0.1 M $MgCl_2$ and immediately centrifuged at 12000 x g to collect the cells. The supernatant was removed and the cell pellet was then resuspended in 2 ml of ice-cold $CaCl_2$. The rest of the procedure was then carried out as described by Sambrook *et al.* (1989D). Transformations carried out using this modified method gave a transformation frequency of 50-200 plaque forming units (pfu)/ng of DNA.

2.2.4.4) Rescreening of Plaques.

Following transformation, 12 small, well-spaced plaques were selected and screened by restriction enzyme analysis using *Asp*718 and *Bam*H1/*Xba*1 and by oligonucleotide hybridisation as previously described. The purpose of this step was twofold. Firstly restriction analysis using *Asp*718 and *Bam*H1/*Xba*1 confirmed that the 328-bp fragment was present in the clone selected and that it had been inserted as a single copy. Secondly

oligonucleotide hybridisation screening served to re-confirm the presence of the mutation and that the fragment was inserted in the correct anti-clockwise orientation. Once a clone was identified as possessing these features, the 328-bp *Asp*718 fragment was sequenced using all three primers as previously described so as to reconfirm the presence of the correct mutation in the clone.

2.2.4.5) Construction of the AR120/pDPTPRO4+pGALKRTE Expression System.

The mutagenized expression was constructed by ligation of the 2.43-kb *Bam*H1/*Xba*1 RT coding region from M13-RT^m to the wild-type 5.82-kb *Bam*H1/*Xba*1 fragment of pGALKRTE to create the 8.25-kb pGALKRTE expression plasmid. Restriction endonuclease digests of both M13-RT and pGALKRTE using *Bam*H1 and *Xba*1 were carried out in a final volume of 13.3 µl consisting of 10 µl of DNA containing solution, 10 units (1 µl) of both *Bam*H1 and *Xba*1 and 1.3 µl and 10 x restriction endonuclease buffer M (as supplied by Boehringer-Mannheim) and incubated at 37°C for 1 hour. The 2.43-kb fragment of M13-RT^m and the 5.82-kb fragment of pGALKRTE were obtained by electroelution from 1% agarose gels as previously described. Ligations were performed as described in Appendix A. The complete ligation reaction was then used to transform competent AR120 cells which already contained the plasmid pDPTPRO4 (AR120/pDPTPRO4) under chloramphenicol selective pressure to form the complete expression vector for HIV-1 RT known as AR120/pDPTPRO4+pGALKRTE^m. The method for the preparation

of competent AR120/pDPTPRO4 was the same as that for JM101 except that chloramphenicol, at a final concentration of 25 µg/ml, was used during the initial growth of the cells. Competent AR120/pDPTPRO4 were transformed to ampicillin resistance with pGALKRTE^{m}. The process of transformation was as previously described for the M13-RT with the following modifications. After heat-shocking of the competent cells in the presence of the ligation mixtures, 1 ml of LB was added to the cells and the mixture incubated at 37°C for 1 hour. The cells were then plated out in 400 µl and 600 µl aliquots onto Luria Agar (LA) plates (see Appendix F) containing chloramphenicol and ampicillin at final concentrations of 25 µg/ml and 50 µg/ml respectively (LA Amp^{50}/Cm^{25}). This is selective for only those AR120/pDPTPRO4 cells that have been transformed with the pGALKRTE^{m} plasmid. The transformation procedures typically produced between 100 and 300 colonies forming units (cfu)/ng of DNA used.

Following construction of the AR120/pDPTPRO4+pGALKRTE^{m} RT expression system, stocks were made of six of the colonies (chosen at random) and restriction analysis was carried out so as to confirm the presence of both plasmids in the clones selected. One clone was then selected for use in subsequent steps.

2.2.5) Expression of Authentic HIV-1 RT Enzyme.

2.2.5.1) Small-scale Induction of HIV-1 RT Expression in AR120/pDPTPRO4+pGALKRTE.

Prior to the large-scale induction of wild-type and mutant RT expression and enzyme purification, a smaller scale version of the procedure was carried out. This involved inoculating the 30 ml of LB Amp^{50}/Cm^{25} with 300 µl of an overnight culture of AR120/pDPTPRO4+pGALKRTE^{m} followed by incubation at 37°C with vigorous aeration. The growth of the culture was monitored spectrophotometrically (using a Phillips Pye Unicam UV8800 UV/Vis spectrophotometer) by periodically measuring its absorbance at a wavelength of 650 nm (A_{650}). When the A_{650} reached a value of between 0.5 and 0.6 (mid-log) aliquots of 1 ml and 7 ml were simultaneously removed. The 1 ml aliquot was used to measure the A_{650} precisely and the 7 ml aliquot was centrifuged to recover the uninduced sample. In addition to this a further 1.5 ml of cells were simultaneously removed for later use for DNA extraction and confirmation of the presence of the two plasmids in the uninduced sample. Nalidixic acid was then added to the remaining 20.5 ml to a final concentration of 60 µg/ml and the cells were incubated at 37°C for a further 5-6 hours with vigorous aeration. After this period, the same three aliquots (1 ml, 7 ml and 1.5 ml) were simultaneously withdrawn from the induced culture and used for A_{650} determination, protein extraction and DNA extraction, as for the uninduced samples. All steps of this procedure were carried out with a parallel culture of AR120/pDPTPRO4+pGALKRTE^{wt}.

Plasmid DNA was isolated from the two 1.5 ml samples (along with the corresponding samples from the wild-type culture) according to the procedure previously described for the isolation of M13-RT RF-DNA. The presence of the two plasmids in both the induced and uninduced samples was confirmed by running 5 µl of each DNA sample on a 1% agarose gel for 1-2 hours alongside the wild-type plasmid DNA. In addition to this, 13.3 µl of *BamH1/Xba1* digests of the plasmid DNA were run on the same gel in order to verify the presence of the 2.43-kb fragment *BamH1/Xba1* in both the induced and uninduced cells. The yield of the plasmid DNA obtained at this stage was also assessed as a means of confirming that the mutations had not altered the plasmid copy number in the final construct.

2.2.5.2) Small-scale RT Activity Screening of Mutants.

In order to make a preliminary assessment of the level of RT activity produced by the mutant enzyme relative to the wild-type enzyme, small-scale RT assays were carried out on crude lysates of the 7 ml samples of induced and uninduced cells.

All volumes described in the procedure are for 7 ml of the induced cell samples. Volumes for the uninduced samples were adjusted proportionally according to the relative A_{650} values of the respective cultures recorded when the cells were removed as previously described. The purpose of this step was to correct for the RT activity contributed by *E.coli* DNA polymerase I.

After pelleting of the 7 ml of cells in a 1.5 ml microfuge tube, the cells were resuspended in 0.8 ml of ice-cold Buffer A [50 mM Tris.HCl (pH 8), 10% glycerol, 0.1% 2-mercaptoethanol]. PMSF, leupeptin and pepstatin were added to final concentrations of 0.18 mg/ml, 1 µg/ml and 1 µg/ml respectively. This was followed by the addition of EDTA and lysozyme to final concentrations of 5 mM and 0.3 mg/ml respectively. The mixture was then incubated on ice for 15 minutes after gentle inversion of the tube. NaCl was added to a final concentration of 0.2 M followed by the addition of 100% Triton X-100 to a final concentration of 1%. The mixture was incubated on ice for a further 30 minutes with occasional inversion of the tube to mix. The NaCl concentration of the mixture was raised to 0.5 M and the contents of the tube were mixed well by inversion. A polyethylene glycol (PEG) precipitate was then formed by the addition, dropwise, of 40 µl of Buffer B [50 mM Tris.HCl (pH8), 10% glycerol, 0.5 M NaCl, 18% PEG 2000, 0.1% 2-mercaptoethanol]. The contents of the tube were mixed well by gentle inversion, and were incubated on ice for 15 minutes with occasional mixing. The microfuge tube was then centrifuged for 30 minutes and the supernatant obtained (PEG supernatant) was removed and saved. This PEG supernatant was used as the crude lysate for small-scale RT activity assays.

PEG supernatant dilutions of 10, 20, 60 and 100 fold were made using 0.5 x RT assay buffer (see Appendix B). RT activity assays were carried out in a final volume of 12.5 µl containing 2.5 µl of the diluted PEG supernatant, 67 mM Tris.HCl (pH 7.9),

67 mM KCl, 7 mM $MgCl_2$, 4 mM DTT, 40 µM [^{3}H]-dTTP (860-2640 dpm/pmol), and 1.3 µM RT substrate (poly(rA).dT_{12-18} ([3'-OH]; 20:1 A:T nucleotide ratio)). The reaction mixtures were incubated at 37°C for 20 minutes and were quenched by the addition of 10 µl of 0.1 M EDTA. The reaction rate was found to be linear under these conditions. After quenching, the complete reaction mixture was spotted onto a 2.5 cm DE81 filter disc (Whatman, Maidstone, UK). The discs were allowed to air-dry and were washed twice in 0.5 M Na_2HPO_4 for 5 minutes followed by a 5 minute wash in 0.3 M ammonium formate (pH 8) and rinsing in 100% ethanol and diethyl ether (10 ml per filter). The filter discs were air-dried and were counted for 1 minute in 20 ml Packard polyethylene scintillation vials (Packard Instrument Co., Illinois, USA) containing 10 ml of Packard Insta-Gel II scintillation fluid (Packard Instrument Co., Illinois, USA). A Beckman LS 6000IC scintillation counter was used for liquid scintillation counting. The results obtained were analysed in two ways. Firstly the filter bound radioactivity obtained from the assay using induced cell lysate was compared to that obtained from the uninduced cell lysate of the same mutant, and secondly, the filter bound radioactivity obtained from the induced and uninduced cell lysates of the mutant were compared to the corresponding wild-type results.

2.2.5.3) Large-scale Induction of HIV-1 RT Expression in AR120/pDPTPRO4+pGALKRTE.

The large scale induction of HIV-1 $RT^{wt/B}$ expression in AR120/pDPTPRO4+pGALKRTE involved the inoculation of 2 l of LB Amp^{50}/Cm^{25} with 20 ml of an overnight culture of the appropriate AR120/pDPTPRO4+pGALKRTE clone (wild-type or mutant). As described for the small-scale induction, the cells were incubated at 37°C with vigorous aeration until the A_{650} reached a value between 0.5 and 0.6. The cells were then induced to express RT by addition of nalidixic acid to a final concentration of 60 µg/ml. The cells were incubated for a further 5-6 hours and collected in 250 ml polyethylene centrifuge tubes (Du Pont Co., Wilmington, USA) by centrifugation at 5000 x g for 10 minutes in a Sorvall RC5C centrifuge using a Sorvall GSA rotor. The pelleted cells (approximately 5 g) were resuspended in 20 ml of Buffer A (as previously described). An aliquot of 1.5 ml of these cells was used to prepare the 2.43-kb *Bam*H1/*Xba*1 RT coding fragment which was recloned into M13-mp11 to recreate M13-RT^{wt} as previously described. Recombinant phages were sequenced through the *Asp*718 fragment region in order to reconfirm the presence of the appropriate sequence in the induced cells to be used for enzyme purification. The remainder of the induced cells were resuspended in Buffer A and were stored at -70°C until required.

2.2.6) Enzyme Purification.

2.2.6.1) Cell Lysis.

All steps in the enzyme purification procedure were carried out at 4°C. The resuspended cells were thawed and the volume increased to 83 ml by the addition of a further 64.5 ml of Buffer A (as previously described). After addition of leupeptin (1 µg/ml), pepstatin (1 µg/ml), PMSF (0.18 mg/ml), EDTA (5 mM) and freshly prepared lysozyme (0.3 mg/ml) the mixture was stirred slowly for 15 minutes in a glass beaker. NaCl was added to a final concentration of 0.2 M followed by the addition of 100% Triton X-100 to a final concentration of 1%. The mixture was stirred for a further 1 hour with occasional mixing using a clean glass rod in order to completely dissolve the Triton X-100. NaCl was added to increase the salt concentration to 0.5 M followed by the addition, over a period of 2 minutes, of 41.7 ml of Buffer B (as previously described). The mixture was stirred for a further 15 minutes before centrifugation for 25 minutes in 4 x 30 ml centrifuge tubes at 12000 x g in a JA20 rotor at 4°C. The supernatant (± 120 ml) was decanted and the volume made up to 360 ml with 2 volumes of Buffer A.

2.2.6.2) Phosphocellulose P11 Column Chromatography.

Phosphocellulose P11 (Whatman, Maidstone, UK), an anionic exchange medium used for the initial enzyme purification step, was prepared as follows. 25 volumes of 0.5 M NaOH was added to

30 g of solid phosphocellulose P11 and the mixture stirred for using a glass rod. The resulting slurry was allowed to settle for 5 minutes and was filtered through a 125 mm S&S filter paper (Whatman, Maidstone, UK) using a Buchner funnel. The paste obtained was washed with 1l of H_2O and refiltered as before. The pH of the filtrate was then read using a pH indicator strip (Merck, Darmstadt, Germany). This cycle was repeated until the pH of the filtrate fell below 10. The phosphocellulose P11 paste was resuspended in 25 volumes of 0.5 M HCl and washed with H_2O as previously described until the pH of the filtrate was greater than 3. The paste was resuspended in 1 litre of 150 mM Tris.HCl (pH 8) and 2 M Tris base was added until the pH was approximately 8. The slurry was then refiltered and the paste resuspended in 1 l of Buffer A and allowed to settle. The pH of the filtrate was determined accurately using a Beckman σ-series pH meter. If the pH was not 8, then the slurry was filtered again and the phosphocellulose P11 resuspended in 1l of fresh Buffer A. The slurry was allowed to settle and any fine particles present in the supernatant were removed. The equilibrated phosphocellulose P11 was stored at 4°C.

Chromatography columns of phosphocellulose P11 for enzyme purification were prepared by loading 100 ml of the phospocellulose P11 slurry into a 280 mm x 28 mm glass chromatography column (Bio-Rad, Richmond, California, USA) and allowing the phosphocellulose to set up overnight. This gave rise to a final phosphocellulose column of approximately 50 ml. The excess fluid was removed from the top of the column and the

phosphocellulose P11 was equilibrated with 250 ml of Buffer C [50 mM Tris.HCl (pH 8), 10% glycerol, 2% PEG 2000, 0.17 M NaCl and 0.1% 2-mercaptoethanol]. This was carried out by passing Buffer C through the column at a flow rate of approximately 1 ml/minute. The 360 ml of cell lysate was loaded onto the column at a flow rate of 0.8-1.5 ml/minute and the column washed with a further 300 ml of Buffer C at a reduced flow rate of 0.3 ml/minute. Following washing, a linear salt gradient of 0.17-0.75 M NaCl (total volume 160 ml) was passed through the column with 3 ml fractions being collected in plastic tubes using a Bio-Rad Model 2110 fraction collector. On completion of the gradient, 20 µl aliquots of fractions 20-39 were analysed by RT activity and electrophoresis on a 10% SDS polyacrylamide gel (SDS-PAGE) (see Appendix C). The peak, cleanest 10 fractions for the 66-kD and 51-kD RT polypeptides were pooled and immediately placed in a dialysis bag and dialysed overnight against 2 l of 50 mM Tris.HCl (pH 8), 10% glycerol, 0.01 M NaCl and 0.1% 2-mercaptoethanol. The dialysis buffer was changed the next morning and the pooled fractions dialysed for a further 2 hours.

2.2.6.3) Q-Sepharose Column Chromatography.

Q-Sepharose, a fast flow anionic exchanger, was obtained from Sigma (Sigma, St.Louis, Missouri, USA) in a pre-swollen form for immediate use in column chromatography. A 15 ml Q-Sepharose column was poured in a 200 mm x 8 mm glass column (Bio-Rad, Richmond, California, USA) and the column equilibrated with 200 ml of 50 mM Tris.HCl (pH 8), 10% glycerol, 0.1% 2-

mercaptoethanol, 0.01 M NaCl which was passed through the column at a flow rate of 1 ml/minute. The ± 30 ml of phosphocellulose P11 pooled fractions were loaded onto the column under at a flow rate of approximately 1 ml/minute. A linear salt gradient of 0.01-0.25 M NaCl (160 ml) was applied to the column at a flow rate of approximately 1 ml/minute and 3 ml fractions were collected as for the phosphocellulose P11 column. On completion of the gradient, 20 µl aliquots of fractions 9-28 were analysed for RT activity and by 10% SDS-PAGE, as for the phosphocellulose P11 fractions, and the 10 cleanest fractions were pooled. Before storage of the Q-Sepharose pooled fractions at -20°C, a 5 ml aliquot of the pool was removed and concentrated using two Millipore Ultra-MC filter units (Millipore, Bedford, Massachusetts, USA) with a molecular weight cut-off of 30-kD using a Millipore Personal Centrifuge until the volume of the aliquot had been reduced to approximately 100 µl.

2.2.6.4) Sephadex G-75 Column Chromatography.

The final stage of the enzyme purification procedure involved passing the concentrated aliquot of the Q-Sepharose pool through a Sephadex G-75 exclusion column. Sephadex G-75 was prepared according to the manufacturers instructions. A 20 ml column was poured in a 300 mm x 5 mm column (Bio-Rad) and equilibrated by passing 200 ml of 50 mM Tris.HCl (pH 8). 10% glycerol, 0.1% 2-mercaptoethanol, 0.1 M NaCl through the column at a flow rate of 0.3 ml/minute. All excess fluid on top of the column was removed and the 100 µl Q-Sepharose pool concentrate

was loaded onto the column. The column was eluted with 50 mM Tris.HCl (pH 8), 10% glycerol, 0.1% 2-mercaptoethanol and 0.1 M NaCl at a flow rate of 0.7 ml/minute and 0.7 ml fractions were collected in 1.5 ml microfuge tubes. 2.5 µl aliquots of fractions 5-15 were tested for RT activity as previously described. The first two fractions showing RT activity were collected so as to avoid the inclusion of contaminating *E.coli* RNase HI. These two fractions were pooled (1.4 ml) and concentrated using Millipore Ultrafree MC filters as previously described until the volume was reduced to approximately 50 µl. Before storage, an equal volume of RT storage buffer [50 mM Tris.HCl (pH 8), 50% glycerol, 0.1% Triton X-100, 1 mM EDTA, 5 mM DTT and 0.1 M NaCl] was added and the purified enzyme was stored at -20°C.

The HIV-1 RT enzyme recovered by this procedure was >95% pure and was free of contaminating *E.coli* RNase HI.

2.2.7) RNase H Assays.

2.2.7.1) Synthesis of the (+)-GAG^{345} RNA transcript.

The RNase H activity of the enzymes was assayed by denaturing gel electrophoretic analysis of the cleavage products of an RNA.DNA oligonucleotide hybrid, (+)-GAG^{345}/RP1. Uniformly labelled (+)-GAG^{345} RNA was prepared by run-off transcription using T7 RNA polymerase. The reaction was carried out in a volume of 40 µl containing 40 mM Tris.HCl (pH 7.5), 12 mM $MgCl_2$, 2 mM spermidine, 10 mM DTT, 1 unit/µl RNasin, 0.25 mM each of rATP,

rCTP, rGTP, 40 µM [α-^{32}P] rUTP (44 Ci/mmol), 0.3 µg/µl of *Hind*III-digested pGEM-ΔGAG-A and 2 units/µl of T7 RNA polymerase. The reaction mixture was incubated at 39°C for 2 hours and the reaction was stopped by the addition of 40 µl of denaturing gel electrophoresis sample loading buffer [80% (v/v) deionised formamide, 0.1% bromophenol blue, 0.1% xylene cyanol, 89 mM Tris, 89 mM boric acid, 2 mM EDTA].

2.2.7.2) Purification of the (+)-GAG^{345} RNA Transcript.

The (+)-GAG^{345} RNA substrate was purified in two steps. Firstly, the complete reaction mixture was loaded onto a 20 cm x 30 cm x 8 mm 6% polyacrylamide gel containing 7 M urea in 1 x TBE buffer (see Appendix B) which was electrophoresed at 400 V for 1 hour. The gel was placed between two sheets of clean, used film and exposed to X-ray film for 30 minutes. The autoradiograph was overlaid on the gel and the band corresponding to the full length, labelled (+)-GAG^{345} RNA was excised using a clean scalpel blade. The RNA was extracted from the polyacrylamide gel slice by agitation in high ionic strength elution buffer [0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, 0.1 % SDS] for 18 hours at 37°C. This elution method was used in order to avoid contamination of the RNA by linear acrylamide which is associated with the crush elution method, since linear acrylamide is an inhibitor of the HIV-1 RT enzyme. The purified (+)-GAG^{345} RNA was precipitated from the elution buffer by the addition of 2.5 volumes of 100% ice-cold ethanol and collected by centrifugation in a Hägar microfuge. The (+)-GAG^{345} RNA pellet was

washed with 70% ethanol, dried and resuspended in 20 µl of TE buffer (pH 8). The concentration of the (+)-GAG^{345} RNA was determined by spotting 1 µl of the resuspended RNA on a DE81 filter, washing the filter as previously described and measuring the radioactivity of the filter in a scintillation counter.

2.2.7.3) Gel Electrophoretic RNase H Assay:

The substrate for RNase H assays was made by hybridisation of the uniformly labelled (+)-GAG^{345} RNA to the DNA oligonucleotide primer, RP1. Hybridisation mixtures (7 µl) containing 40 mM Tris.HCl (pH 7.9), 40 mM KCl, 1 mM $MgCl_2$, uniformly labelled (+)-GAG^{345} RNA (30 nM, 3 x [illegible] cpm) and 0.5 µM RP1 were placed in a 500 ml water bath at 50°C with hybridisation occurring over a period of ~1 hour by allowing the water bath to cool to room temperature.

RNase H assays of the enzymes were carried out in a final volume of 20 µl consisting of 50 mM Tris.HCl (pH 7.9), 7 mM $MgCl_2$, 5 mM DTT, the ^{32}P-labelled RNA.DNA hybrid substrate as above and either HIV-1$^{wt/m}$ (12.5 nM) or *E.coli* RNase HI (0.4 nM) and incubated at 37°C. Aliquots (4 µl) of the reaction mixture were removed after 0, 5, 15 and 30 minutes and the reactions stopped by vigorous mixing with an equal volume of denaturing gel sample loading buffer (as described previously). The samples (8 µl) were loaded directly onto a 20 cm x 40 cm x 0.4 mm 8% denaturing polyacrylamide sequencing gel and electrophoresed for 2.5 hours at 1500 V. The gel was then covered with clingwrap and

exposed to X-ray film at -70°C using one intensifier screen.

2.2.8) Steady-state Kinetic Studies of Mutant and Wild-type Enzymes.

The method used for the steady-state kinetic studies carried out on the wild-type and mutant enzymes was based on the standard DE81 filter assay for RT activity using a poly(rA).dT_{12-18} template-primer as described in section 2.2.5.2. Reaction mixtures (50 µl) containing 67 mM Tris.HCl (pH 7.9), 57 mM KCl, 20 mM NaCl, 7 mM $MgCl_2$, 0.2 mM EDTA, 50 µM dTTP, 40 µM [^{3}H]-TTP (880-2640 cpm/pmole), 2-200 nM poly(rA).dT_{12-18} RT substrate and 0.5-2 nM RT enzyme were incubated at 37°C. During incubation 12 µl aliquots were removed at 0, 20, 40, 60, 90, 120 and 180 seconds and quenched with 10 µl of 0.1 M EDTA. The quenched reaction mixtures were spotted onto DE81 filters and the filters washed, dried and counted in a scintillation counter as previously described. The results obtained were interpreted using Enz-fitter, a software program for the interpretation of kinetic data (Enz-fitter (version 1.03), Elsevier-Biosoft, Cambridge, UK using a Ghostech 386SX-20C IBM compatible computer).

2.2.9) Thermal Inactivation Studies of the D443A and Wild-type RT Enzymes.

Thermal inactivation studies of the mutant and wild-type enzymes were modelled on those described by Tisdale *et al.* (1991). Prior to the examination of the thermal stability, the

purified enzymes were diluted so that the rate of incorporation of [^{3}H]-TTP was linear over an assay period of 15 minutes as measured by the standard DE81 filter assay for RT activity (see section 2.2.5.2). The standard RT assay mixture was divided into two separate solutions. Solution 1 consisted of 97.8 mM Tris.HCl (pH 7.9), 97.8 mM KCl, 11.7 mM $MgCl_2$ and 6.8 mM DTT. Solution 2 consisted of 4.8 µM poly(rA).dT_{12-18} and 148 µM [^{3}H]-dTTP (3250-9750 cpm/nmol). For each enzyme, duplicate microfuge tubes containing 1.5-3.0 nm diluted enzyme and 51.1 µl of solution 1 were preheated at 37°C for 5 minutes. The tubes were placed in an Eppendorf thermocycler (Eppendorf, USA) set at 45°C and 9.8 µl aliquots were removed after 0, 5, 10, 20, 30 and 40 minutes. The aliquots were immediately placed in a microfuge tube containing 2.7 µl of solution 2 and incubated at 37°C. The reaction was quenched after 15 minutes by the addition of 10 µl of 0.1 M EDTA and the quenched mixture was placed on a DE81 filter disc. The disc was then washed, dried and counted as previously described in section 2.2.5.2 and the results expressed as a percentage of the activity obtained at the 0 time point.

3.) RESULTS.

3.1) Construction of Mutants.

The scheme used for the construction of the expression system for the mutant enzymes is summarised in Figure 9. The purpose of the various steps involved in the construction of the mutants was threefold:

1.) To introduce site-specific mutations within the RNase H region of the RT gene which had been cloned into M13mp11.

2.) To ensure that the correct mutation was present in the clone and that no undefined, second-site mutations had occurred during the mutagenesis procedure.

3.) To transfer the mutated RT gene from M13 into the expression vector pGALKRTE so as to allow large scale production of the mutant enzyme.

Figures 10 to 16 indicate, in more detail, the various stages of this process. Figures 10A and 10B shows the results obtained after the initial screening of 24 plaques from the D443A and D443A/D498N mutants respectively (Figure 9, step 1), directly following oligonucleotide-directed mutagenesis. Following screening, a plaque showing positive hybridisation with the appropriate mutagenic oligonucleotide was chosen in each case, and the presence of the correct mutant sequence in each clone was

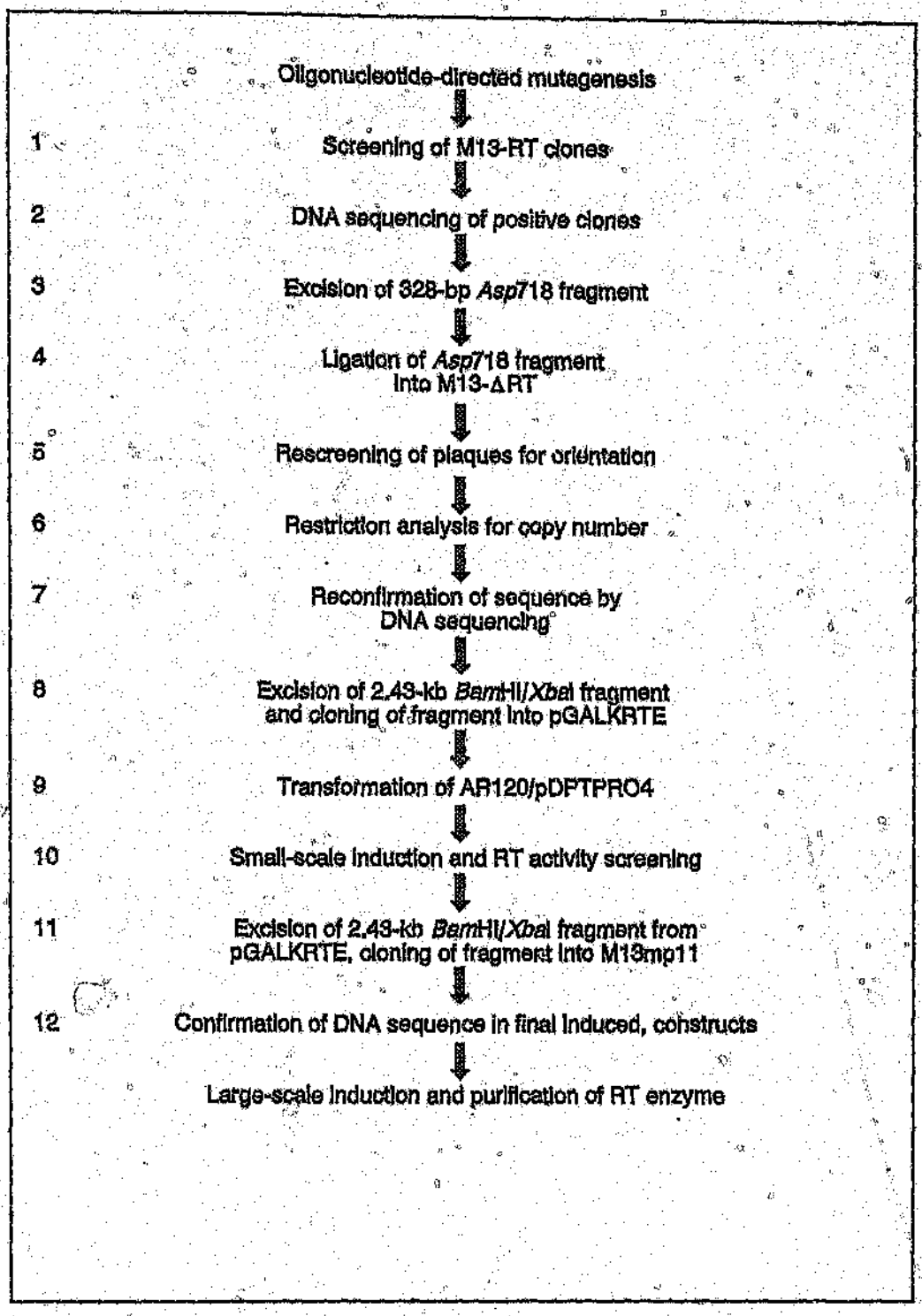


Figure 9: Summary of the scheme used for the construction of the AR120/pDPTPRO4+pGALKRTE RT expression system.

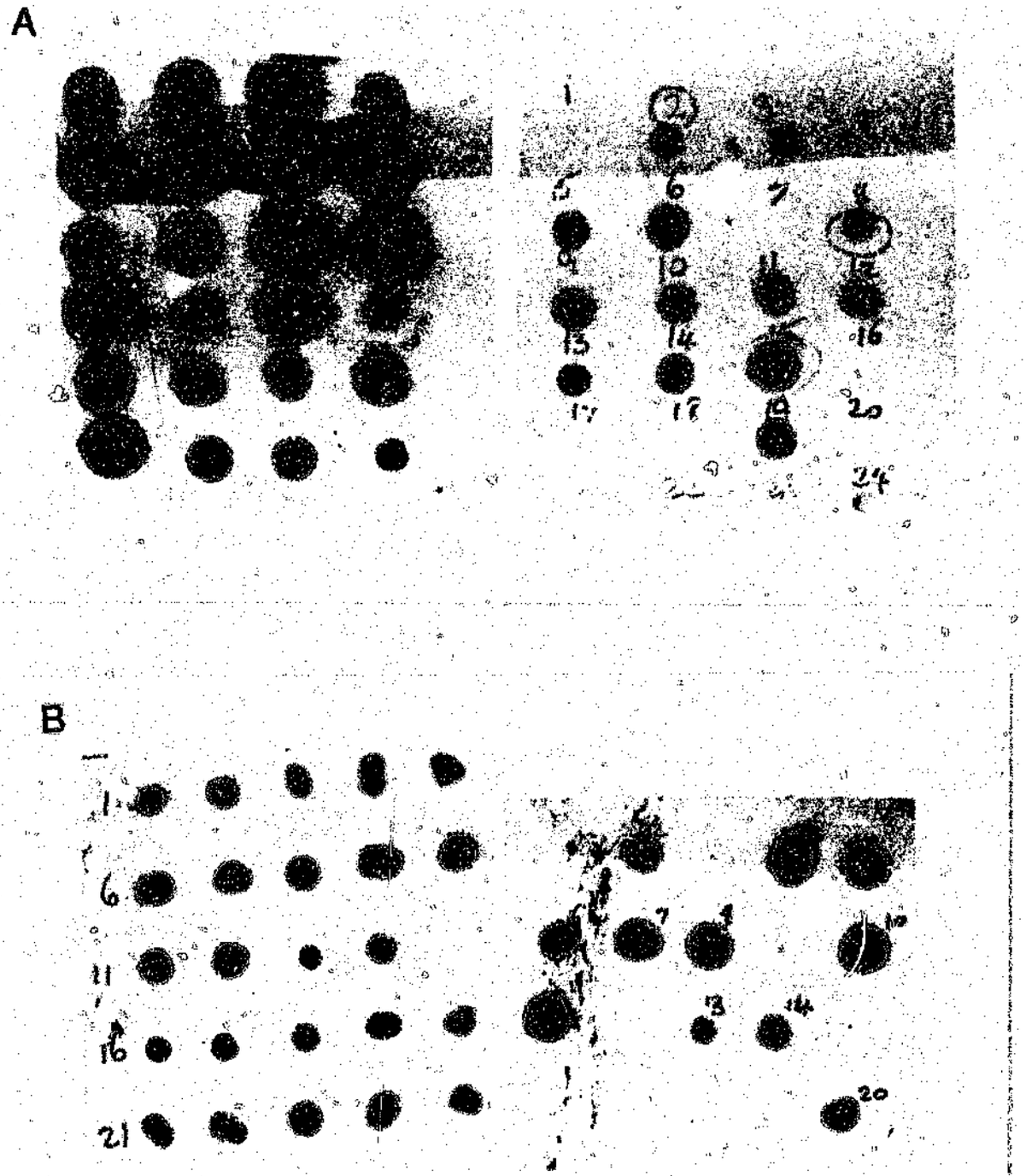


Figure 10: Oligonucleotide screening of 24 clones obtained immediately after the mutagenesis procedure. 10A shows autoradiographs of the D443A mutant clones obtained after washing of the membrane at room temperature and at 64°C. 10B shows autoradiographs of the D443A/D498N mutant clones obtained after washing of the membrane at room temperature and 48°C.

confirmed by DNA sequencing (Figure 9, step 2).

Figure 11 shows the 328-bp fragment obtained after digestion by the restriction endonuclease *Asp*718 of RF-DNA obtained from these plaques (Figure 9, step 3). This fragment was ligated to the vector M13-ΔRT so as to reconstruct M13-RT^m (Figure 9, step 4). Figures 12 shows the results obtained after a second round of oligonucleotide screening of 12 plaques obtained after this step from the D443A and D443A/D498N mutants. (Figure 9, step 5). Positive plaques indicated that the 328-bp fragment had been inserted into the M13-ΔRT vector in the correct anticlockwise orientation. In order to check that the 328-bp fragment had been inserted as a single copy as opposed to a multimer, RF-DNA from the selected positive clones was digested with the restriction endonucleases *Bam*H1 and *Xba*1 so as to obtain the 2.43-kb RT fragment. The digest was analysed on a 1% agarose gel and the size of the fragments examined for insertions (Figure 9, step 6). Figure 13 illustrates a result obtained for the D443A mutant after the restriction analysis of 4 plaques for the presence of a single copy of the 328-bp fragment. The plaque represented by lane 4 shows the presence of a larger *Bam*H1/*Xba*1 fragment than that obtained for the standard used in lane 3 and suggested that two copies of the 328-bp *Asp*718 fragment had been ligated to the M13-ΔRT vector in this case. The plaque represented by the digest in lane 7, which contained a *Bam*H1/*Xba*1 fragment of the correct size (2.43-kb), was used for later steps in the construction of the D443A mutant enzyme. Having reconstructed the M13-RT^m vector and screened for correct

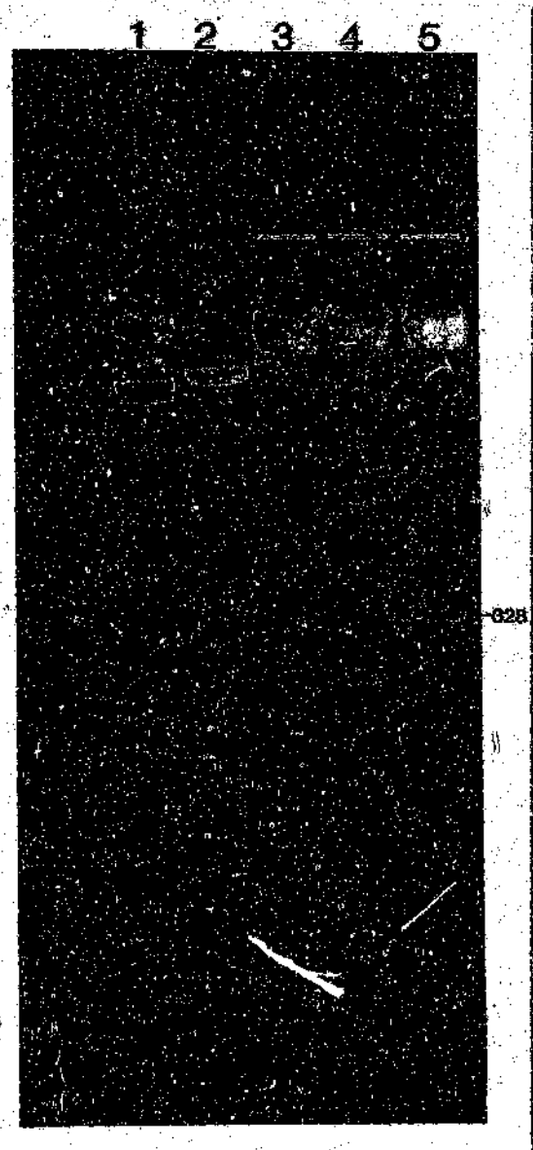


Figure 11: A 1% agarose gel showing the 328 bp obtained after digestion of M13-RT RF-DNA with *Asp*718 (lanes 3-5). Lane 1= MW marker III, Lane 2= undigested RF-DNA.

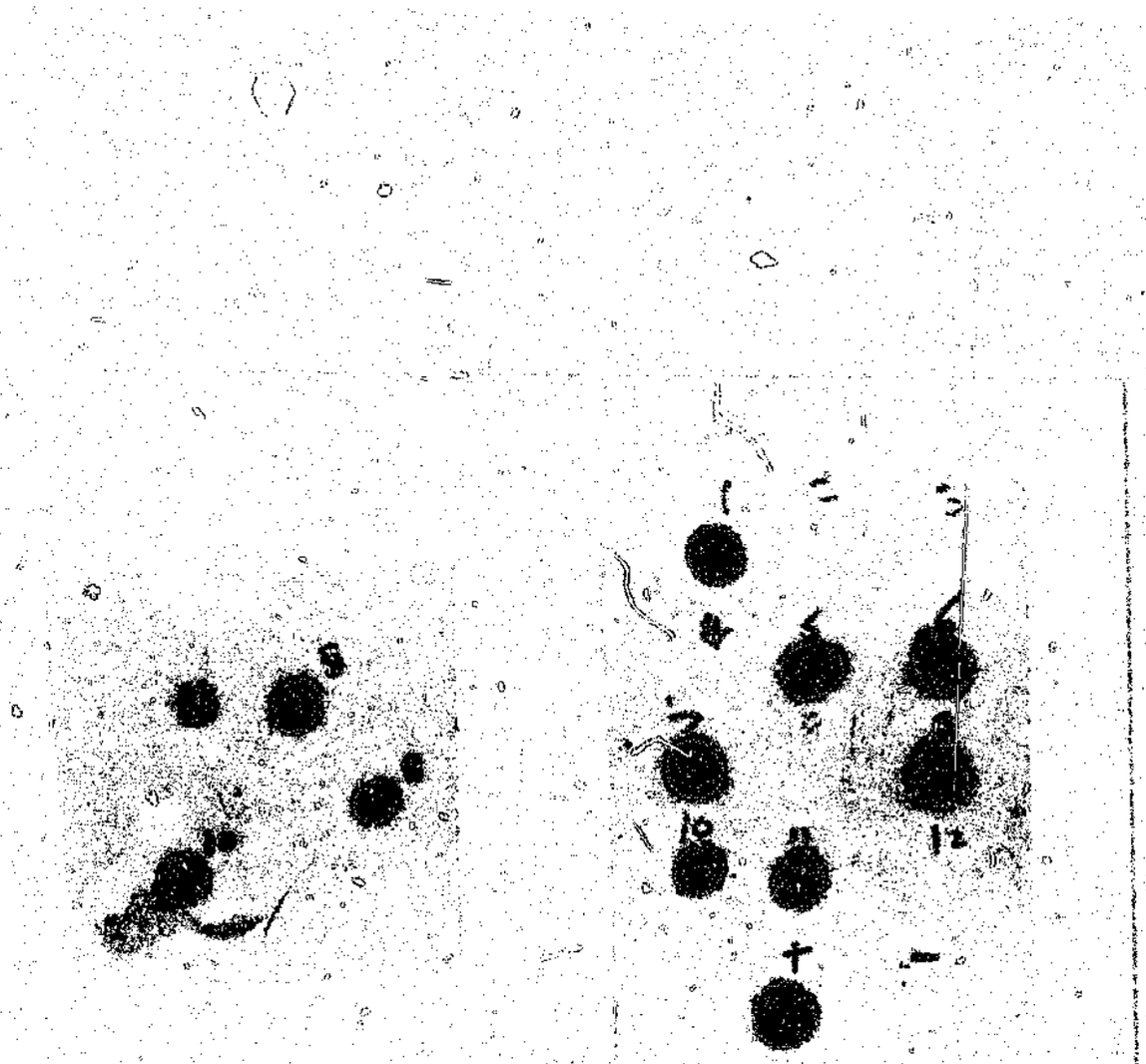


Figure 12: Autoradiographs obtained after oligonucleotide screening of 12 plaques from M13-RT^{D443A} (left) and M13-RT$^{D443A/D498N}$ (right) obtained after ligation of the *Asp*718 fragment to the M13-ΔRT vector.

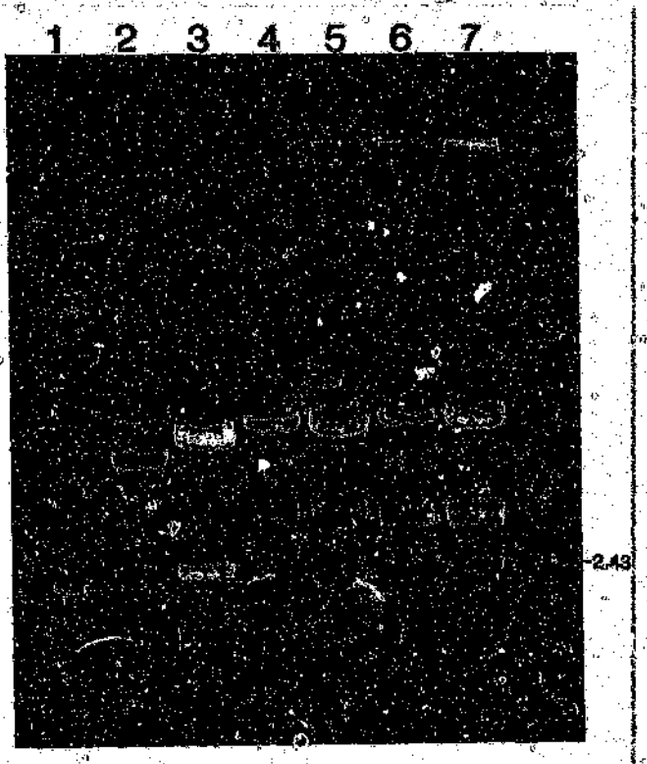


Figure 13: A 1% agarose gel of a *Bam*H1/*Xba*1 digestion of the RF-DNA obtained from the four positive clones of the D443A mutant (lanes 4-7) as illustrated in Figure 12. Lane 1= DNA MW marker III, Lane 2= Undigested pGALKRTE plasmid, Lane 3= *Bam*H1/*Xba*1 digested pGALKRTE plasmid.

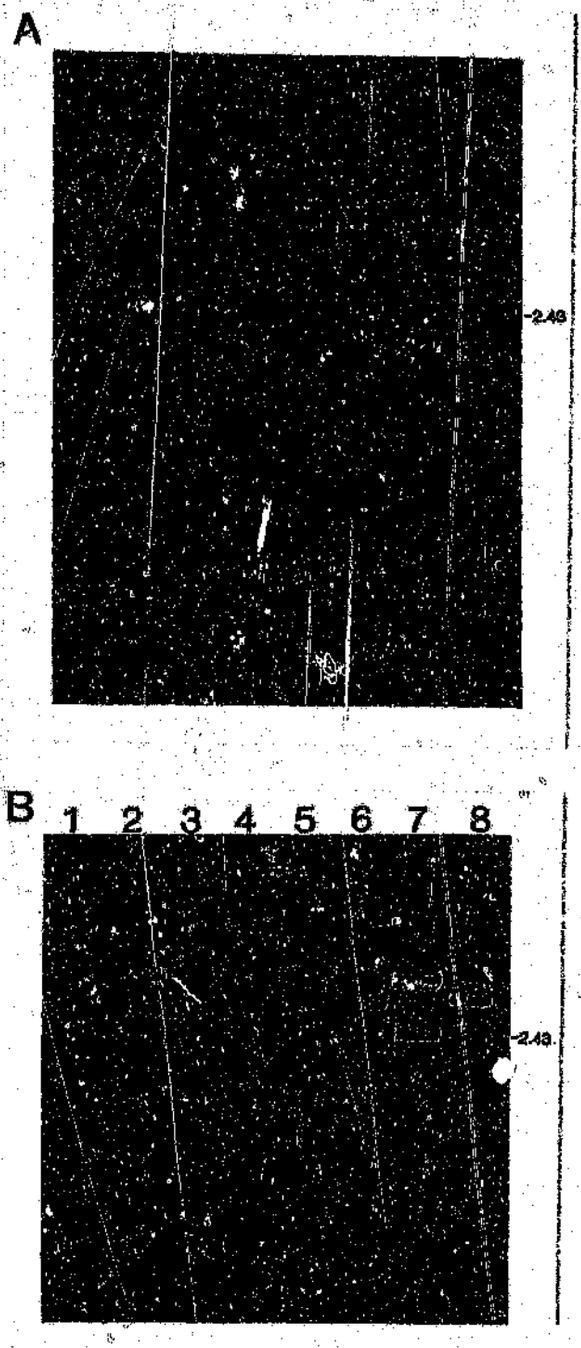


Figure 14: 1% agarose gels of *Bam*H1/*Xba*1 digested RF-DNA obtained from the final M13-RT clones. Figure 14A shows digested D443A RF-DNA (lanes 2-8), DNA MW marker III (lane 1) and undigested D443A RF-DNA (lane 2). Figure 14B shows *Bam*H1/*Xba*1 digested D443A/D498N RF-DNA (lanes 2-5), undigested D443A/D498N RF-DNA (lane 1), *Bam*H1/*Xba*1 digested pGALKRTE plasmid (lane 7) and undigested pGALKRTE plasmid (lane 8).

orientation and copy number, the presence of the correct mutant sequence was reconfirmed by DNA sequencing of the *Asp*718 region (Figure 9, step 7). The overall purpose of steps 3 to 7 was to obviate the need to sequence the entire 2.43-kb RT coding fragment to check for second-site mutations that may have occurred elsewhere in the RT coding region during the mutagenesis procedure.

The entire 2.43-kb RT coding region was then cloned into the RT expression vector pGALKRTE (Figure 9, step 8). Figures 14A and 14B illustrate the results obtained after *Bam*H1/*Xba*1 digestion of the RF-DNA obtained from the final M13-RT^{D443A} and M13-RT$^{D443A/D498N}$ clones prior to electroelution of the 2.43-kb RT fragments and ligation of this fragment to pGALKRTE.

Following construction of the pGALKRTE^{m} expression plasmid and transformation of AR120/pDPTPRO4 with this plasmid, small-scale RT activity screening of the mutants was performed (Figure 9, step 10). Table 2 shows that no RT activity above that of the background contributed by *E.coli* DNA polymerase I could be detected in the crude cell lysates of induced AR120/pDPTPRO4+pGALKRTE$^{D443A/D498N}$. In addition to this, total plasmid DNA was isolated from the induced cells used in this step and digested with *Bam*H1/*Xba*1 so as to check for the presence of the 2.43-kb RT coding fragment in the final construct. Figure 15 shows the results obtained at this step for the wild-type, D443A and D443A/D498N pGALKRTE plasmids. In the case of both the D443A and D498N mutants, the entire 2.43-kb RT fragment isolated from

Dilution factor of crude cell lysates.	RT activity of the crude cell lysate of the induced D443A/D498N RT mutant (cpm).	RT activity of the crude cell lysate of the induced wild-type RT enzyme (cpm).
10	2686	72922
20	1790	61146
60	558	21683
100	264	14187

Table 2: The RT activity of the crude cell lysate of induced AR120/pDPTPRO4+pGALKRTE$^{D443A/D498N}$ versus that of AR120/pDPTPRO4+pGALKRTE^{WT}.

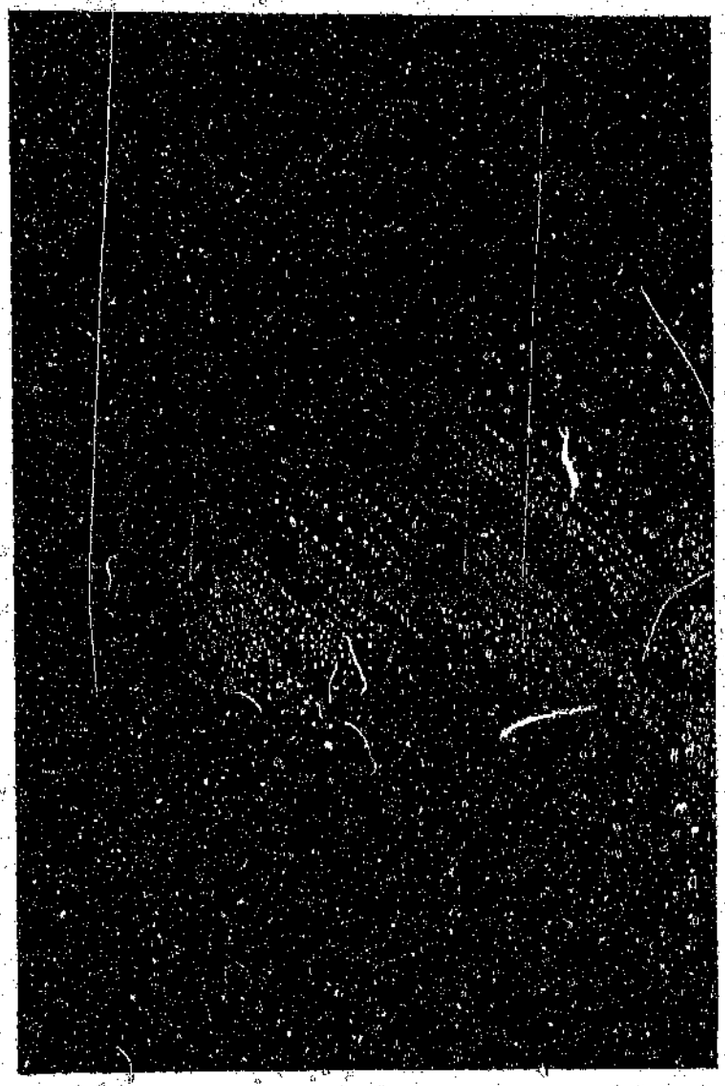


Figure 15: A 1% agarose gel of *Bam*H1/*Xba*1 digests of plasmid DNA obtained from induced cells immediately prior to cell lysis and enzyme purification. Lane 1= DNA MW marker VI, lane 2= *Bam*H1/*Xba*1 digested pGALKRTE^{WT} plasmid, lane 3= digested WT induced plasmid DNA, lane 4= digested D443A induced plasmid DNA, lane 5= digested D443A/D498N induced plasmid DNA and lanes 6-8= undigested samples of lanes 3-5.

the induced cell pellet was cloned into M13mp11 so as to reconstruct M13-RT^m (Figure 9, step 11). Single-stranded DNA was isolated from these clones and the presence of the correct mutant sequence in the final constructs used for expression of the mutant RT enzymes was confirmed by DNA sequencing of the entire *Asp*718 region (Figure 9, step 12). Figures 16 A,B show the appropriate DNA sequences for the D443A and D443A/D498N clones respectively as obtained at this stage immediately prior to large scale expression and purification of the enzymes. The D443A mutant represented by Figure 16A is indicated by the change of the wild-type codon GAT (aspartic acid) to GCA (alanine) at nucleotides 1522 to 1524 of the *Bam*H1/*Xba*1 RT coding fragment. The D443A/D498N mutant represented by Figure 16B is indicated by two mutations, the D443A mutation as previously described and the D498N mutation which indicated by the change of the wild-type codon GAC (aspartic acid) to AAC (asparagine) at nucleotides 1687 to 1689 of the *Bam*H1/*Xba*1 fragment.

3.2) Analysis of Mutants.

3.2.1) Protein Purification.

The purification of the wild-type and mutant enzymes, after large-scale induction of enzyme expression, was carried out in three stages: (1) phosphocellulose P11 column chromatography (2) Q-Sepharose column chromatography, and (3) Sephadex G-75 column

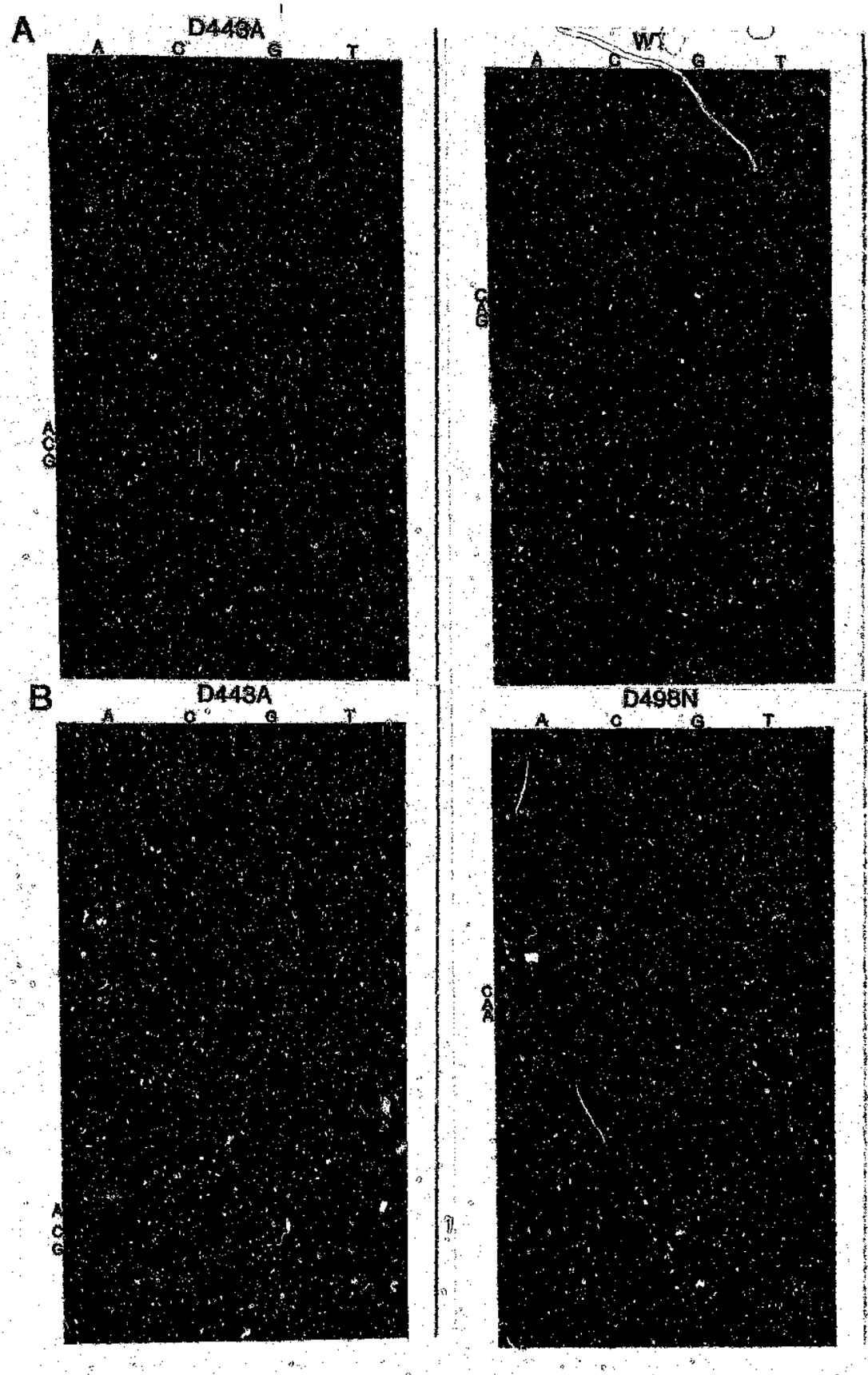


Figure 16: The DNA sequence at the 443 and 498 positions of the D443A (16A) and D443A/D498N (16B) RT enzymes as obtained from the final induced cells.

chromatography respectively. An overall scheme for the purification process is illustrated in Figure 17.

Figure 18 shows the results of the first stage of the purification of the D443A mutant. The photograph shows a 10% SDS-PAGE gel of 20 µl aliquots of fractions 20-39 which were obtained after the cell lysate had been passed through a column containing phosphocellulose P11. The ten most pure fractions (fractions 25 to 34) were collected, dialysed and passed through a Q-Sepharose column.

Figure 19 shows a photograph of a 10% SDS-PAGE gel of 20 µl aliquots of fractions 9-28 obtained after Q-Sepharose column chromatography of the D443A mutant enzyme. The ten most pure fractions (fractions 15 to 24) were pooled and a 5 ml aliquot of the pool was concentrated and applied to the Sephadex G-75 column. The RT assay results from the analysis of 2.5 µl aliquots of the first 11 fractions collected from the Sephadex G-75 column during the purification of the D443A mutant enzyme are shown in Table 3. The first two fractions showing substantial RT activity, fractions 7 and 8, were pooled, concentrated and used for analysis of the properties of the mutant enzyme.

Figure 20 shows a photograph of the fractions collected from the phosphocellulose P11 column during the first step of the purification of the D443A/D498N mutant. The result illustrated that no protein corresponding to the 51 and 66-kD subunits of the

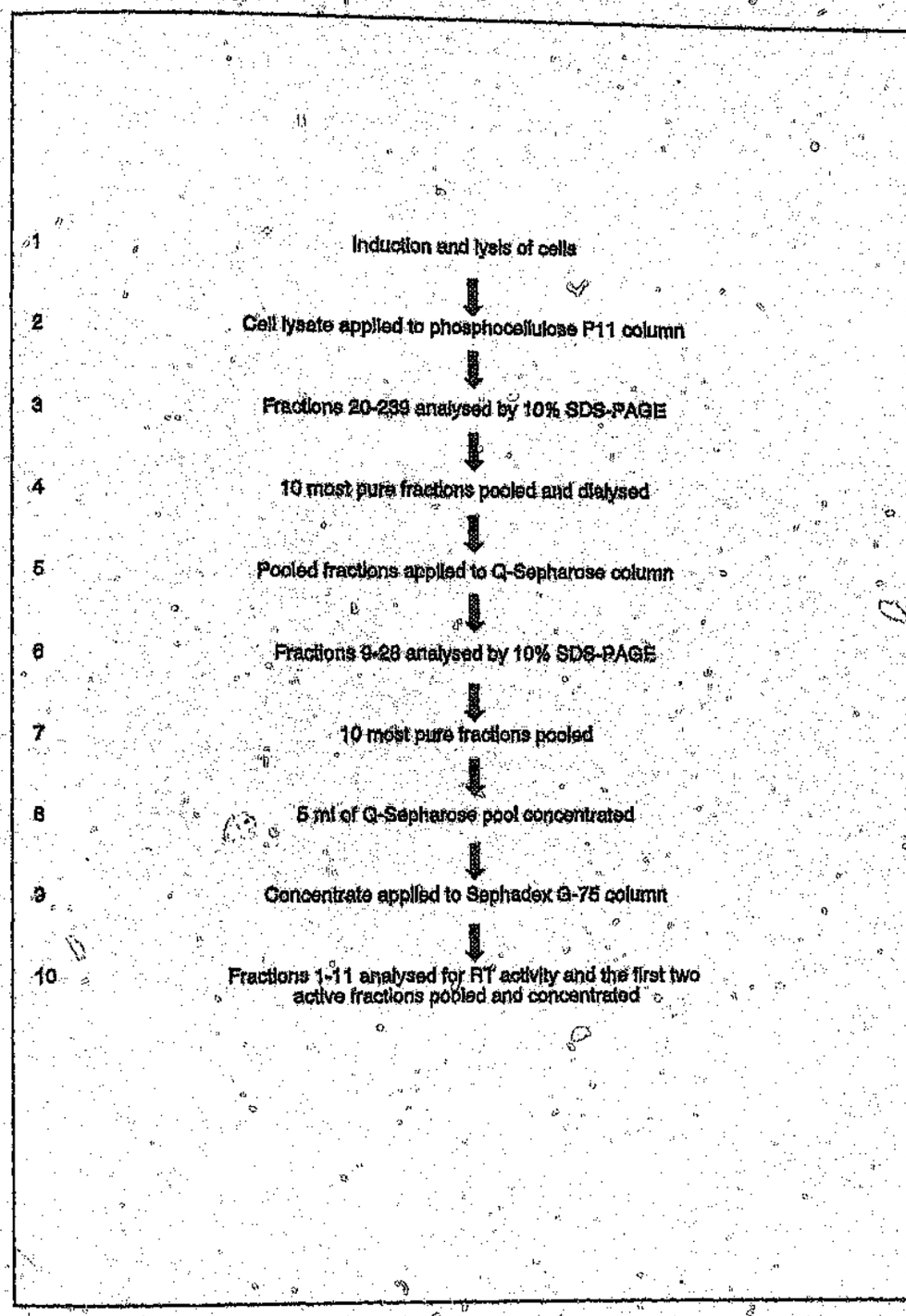


Figure 17: Summary of the procedure used for the purification of authentic HIV-1 RT.

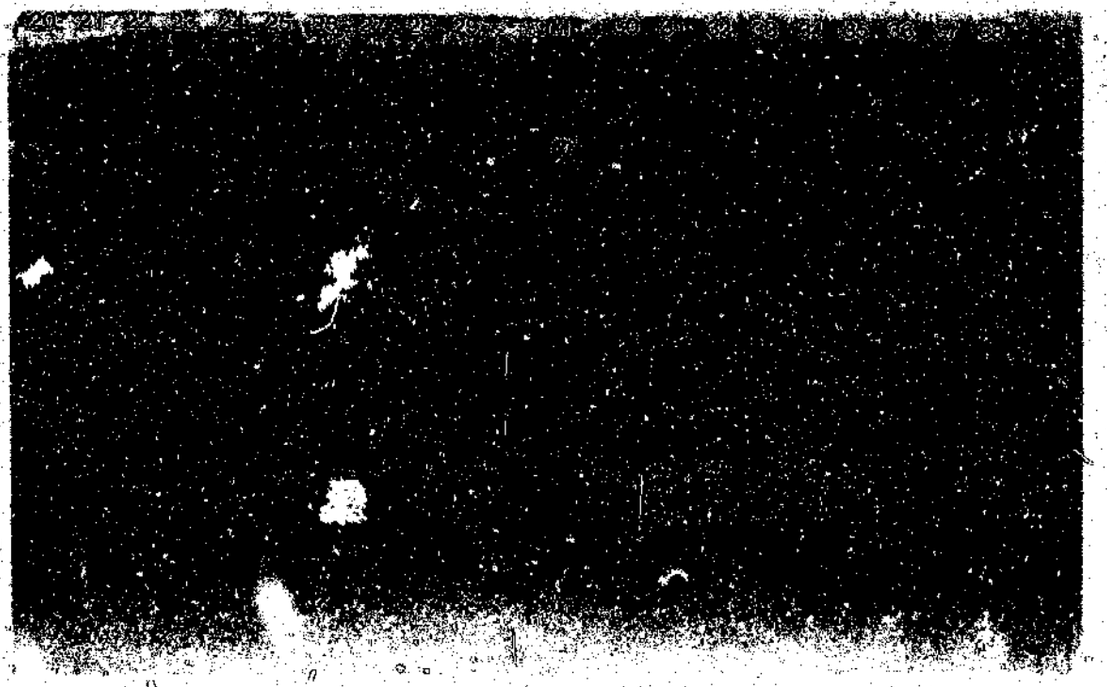


Figure 18: A 10% SDS polyacrylamide gel of 20µl samples of fractions 20 to 38 as obtained after phosphocellulose P11 column chromatography of the D443A cell lysate. Molecular weight markers are indicated (M).



Figure 19: A 10% SDS polyacrylamide gel of 20 µl samples of fractions 9-28 as obtained after Q-Sepharose column chromatography of the D443A enzyme. A molecular weight standard indicated (M).

RT enzyme was present in the fractions obtained from the phosphocellulose P11 column. The low expression level, lack of RT activity and apparent instability of this enzyme prevented all further attempts to purify this particular mutant.

Figure 21 shows a 10% SDS-PAGE gel of the final product of the purification procedure. Lane 1 is the molecular weight standard as indicated, lanes 2 and 3 are the purified wild-type enzyme and the purified D443A mutant enzyme respectively. The result illustrates that both the wild-type and D443A enzymes were isolated as an equimolar p66/p51 polypeptide mixture and that no distinction could be made between the two enzymes at this level.

G-75 fraction numbers.	RT activity of fraction (cpm).
1	45
2	58
3	60
4	53
5	67
6	219
7	138843
8	136502
9	145754
10	146207
11	149956

Table 3: RT activity of 2.5 µl samples of the first eleven fractions obtained after Sephadex G-75 column chromatography.

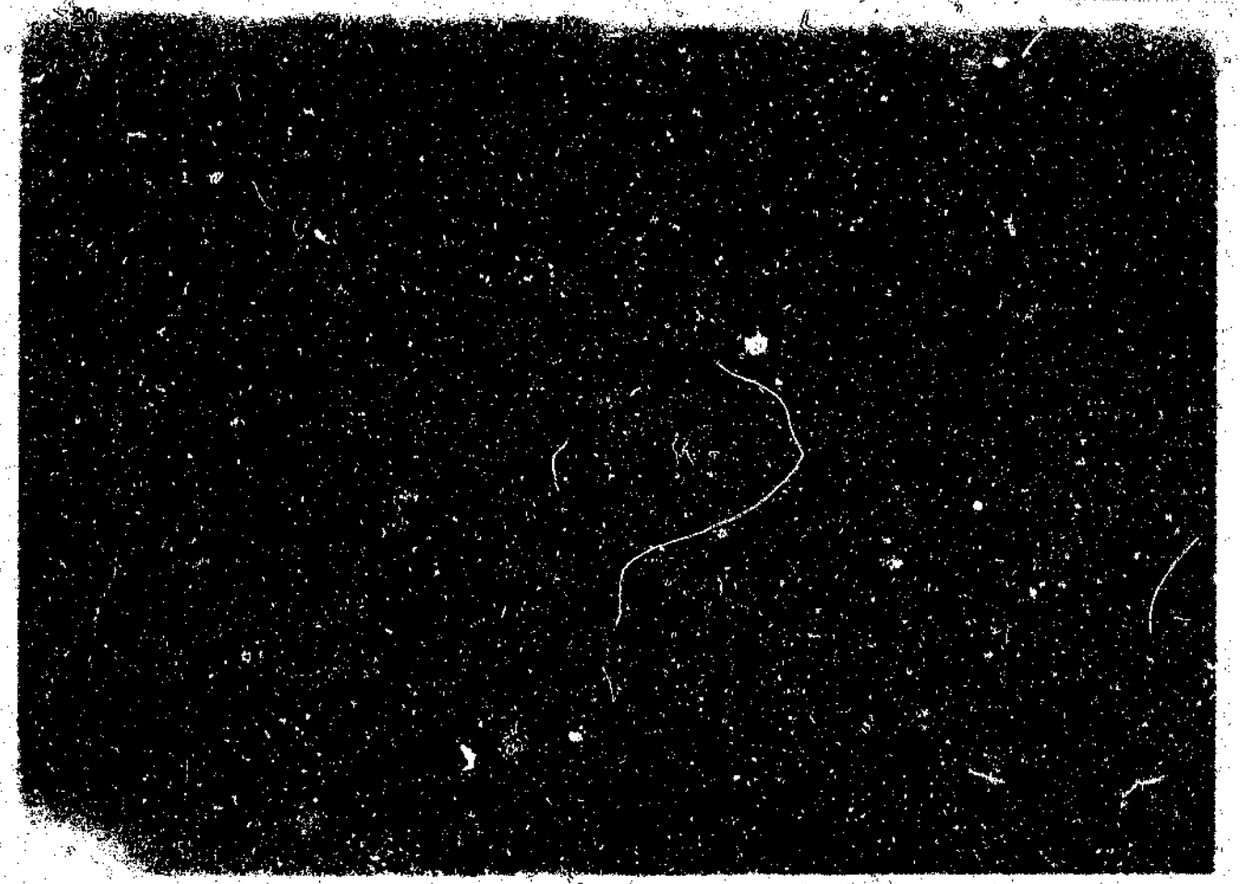


Figure 20: A 10% SDS acrylamide gel 20 µl samples of fractions 20 to 38 as obtained after phosphocellulose P11 column chromatography of the D443A/D498N mutant enzyme. A purified RT standard is indicated (M).

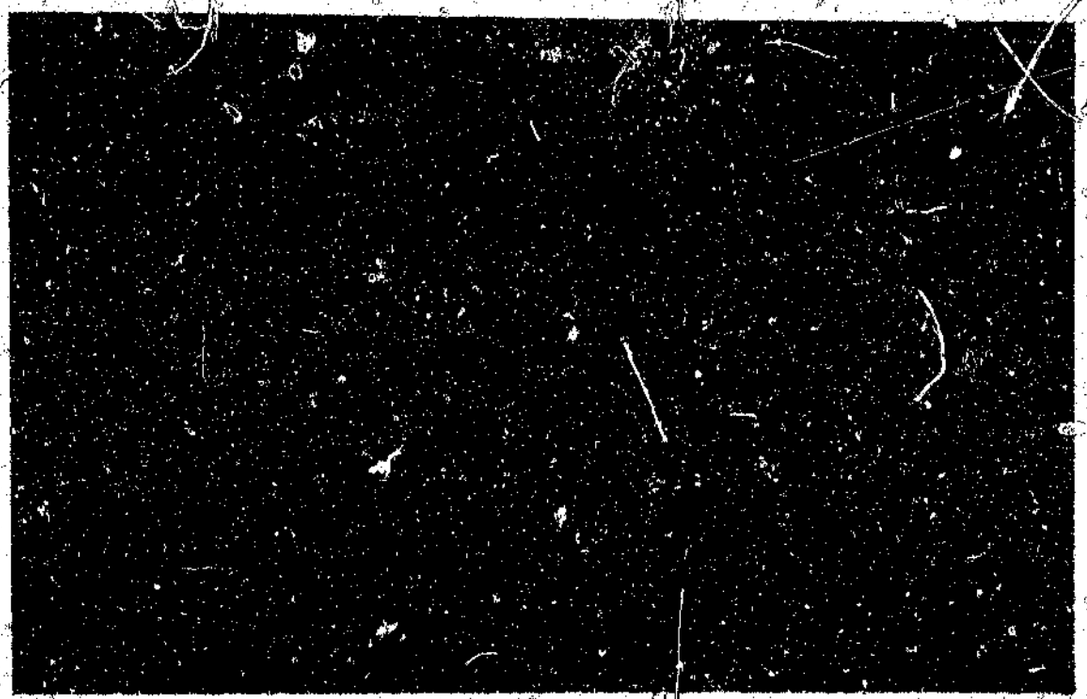


Figure 21: A 10% SDS polyacrylamide gel of the final product of the purification procedure. Lane 1= molecular weight standard, lane 2= purified wild-type RT enzyme, lane 3= purified D443A RT enzyme.

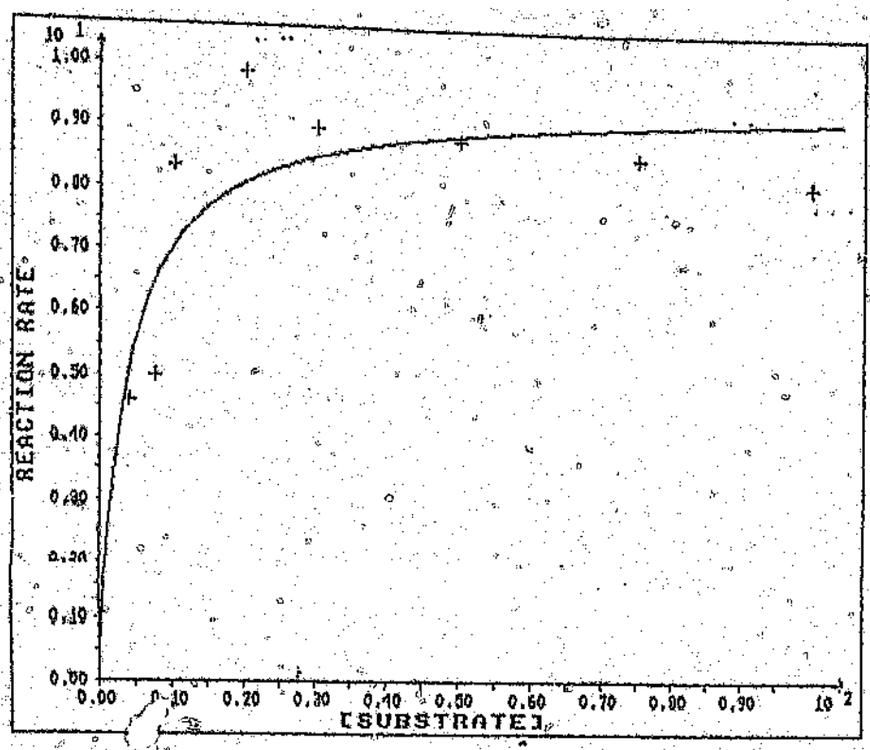
3.2.2) Steady-state Kinetic Analysis of the RT activity of the D443A mutant.

Comparative steady-state kinetic analysis of the RT activity of the D443A RT mutant and the wild-type enzyme was carried out using the homopolymer RT substrate poly(rA).dT_{12-18} as described in "Materials and Methods". The initial rates measured were directly proportional to the concentration of the RT enzyme over the range of substrate concentrations used (2-200 nm). At each substrate concentration, product concentrations were measured in duplicate at each time point (0, 20, 40, 60, 90, 120, 180 seconds) from which initial rates were calculated. Figures 22A and B show a plot of the substrate concentration versus reaction velocity for the wild-type and D443A enzymes respectively. Using the software program EnzFitter, the K_m and K_{cat} kinetic parameters were calculated from the data illustrated in Figures 22A and B. The comparative values obtained for these constants are summarised in Table 4. These data indicated that the K_m and K_{cat} values obtained for polymerisation by the D443A mutant enzyme using a homopolymer substrate were not significantly different from that of the wild-type enzyme.

3.2.3) Thermal Inactivation Studies of the D443A and Wild-Type RT Enzymes.

An assessment of the thermal stability of the D443A mutant enzyme relative to the wild-type enzyme was made by exposing the enzymes to a constant temperature of 45°C for 0-40 minutes.

A



B

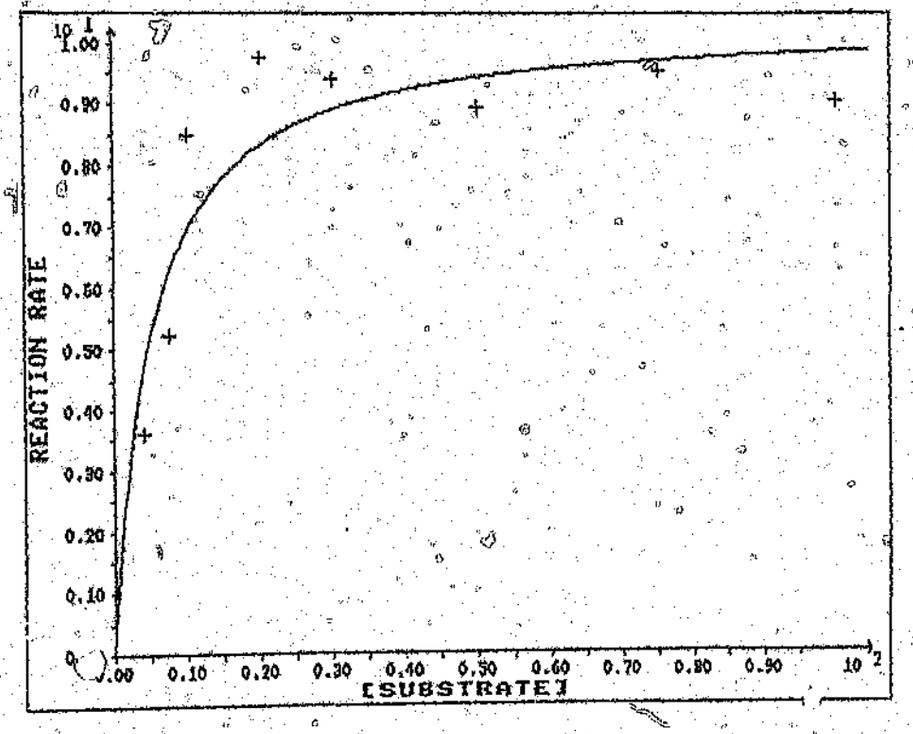


Figure 22: Reaction velocity versus substrate concentration (nM) for RNA-dependent DNA polymerisation using a poly(rA).dT_{12-18} substrate for the wild-type (A) and the D443A (B) HIV-1 RT enzymes.

Enzymes	K_m (nM)	K_{cat} (s^{-1})
D443A	4.7 ± 1.5	5.3 ± 1.3
Wild-type	3.5 ± 0.6	4.5 ± 0.4

Table 4: Kinetic constants obtained for RNA-dependent DNA polymerisation by the D443A and wild-type HIV-1 RT enzymes using a poly(rA).dT_{12-18} homopolymer substrate.

Duplicate aliquots were removed after 0, 5, 10, 20, 30 and 40 minutes and were assayed for RT activity. The activity at each time point was calculated as a percentage of the initial RT activity recorded at the 0 time point. The results obtained for the rates of thermal inactivation of the wild-type and D443A RT enzymes are shown in Figure 23. These results suggest that the D443A mutation has no significant effect on the thermal stability of the mutant enzyme relative to that of the wild-type with a 50% reduction in activity of both enzymes occurring after approximately 35 minutes at this temperature.

3.2.4) Effect the D443A Mutation on the RNase H activity.

The time courses for RNase H cleavage of the (+)-GAG^{345}/RP1 RNA.DNA hybrid catalysed by *E.coli* RNase HI, the wild-type RT enzyme and the D443A mutant RT enzyme are illustrated in Figure 24. In the case of both *E.coli* RNase HI and the wild-type RT

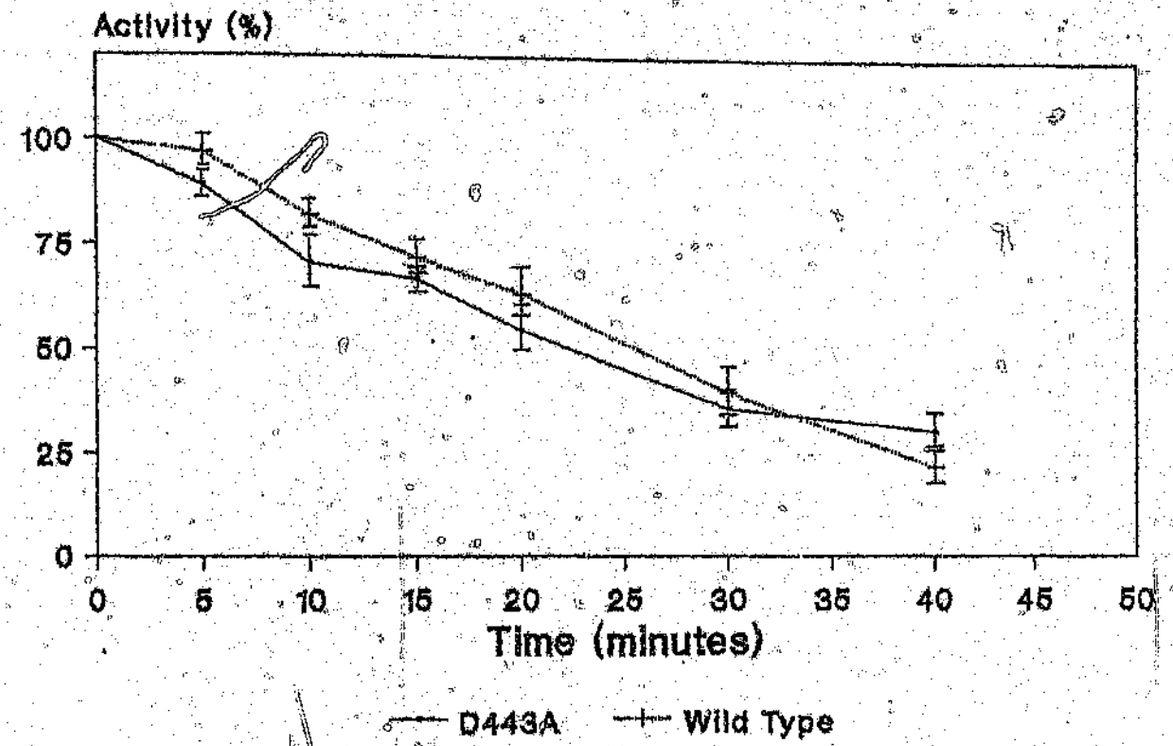


Figure 23: The percentage decrease of RT activity versus time after exposure of the D443A and the wild-type RT enzymes to a temperature of 45°C.

enzyme, complete degradation of the (+)-GAG^{345} RNA to a variety of smaller fragments occurs within 30 minutes. The difference in fragment distribution produced by *E.coli* RNase HI (panel A) and the wild-type RT enzyme (panel B) suggests that the wild-type enzyme is free of *E.coli* RNase HI contamination. The lack of any detectable degradation seen in the case of the D443A mutant enzyme (panel C) suggests that this enzyme lacks an RNase H activity.

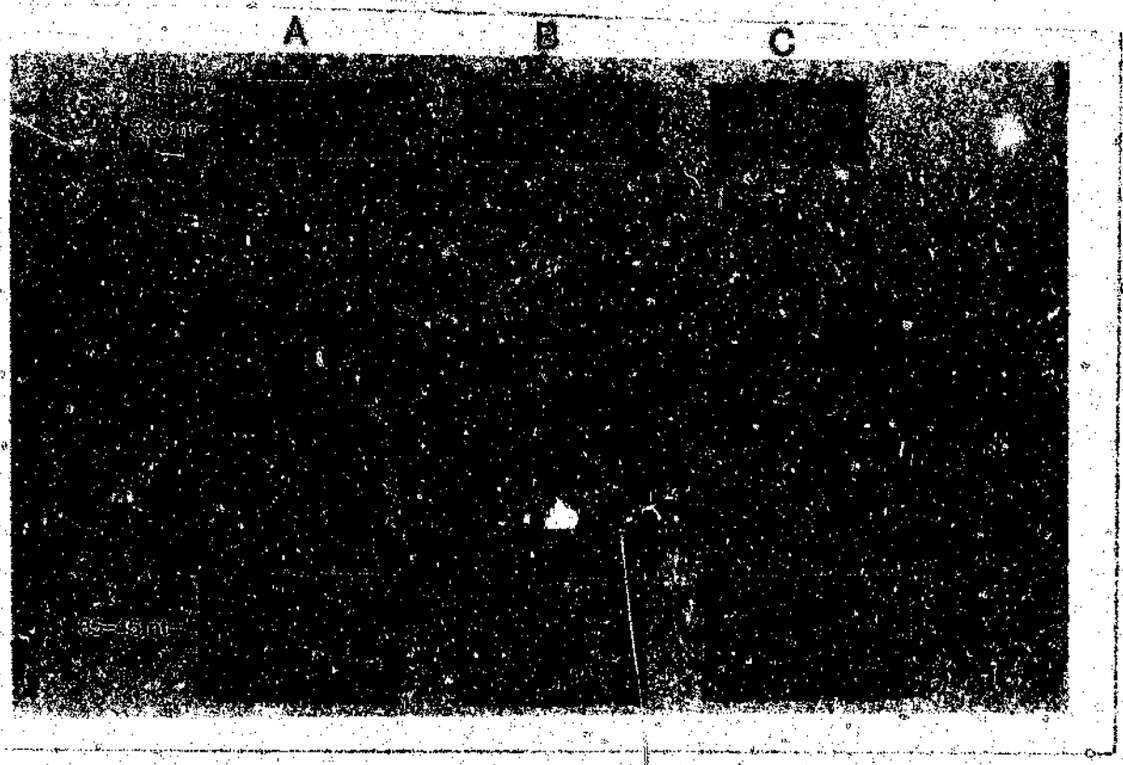


Figure 24: RNase H mediated cleavage of (+)-GAG^{345} RNA by *E.coli* RNase HI (panel A), wild-type RT (panel B) and D443A RT (panel C).

4) DISCUSSION AND CONCLUSION.

The roles of the highly conserved aspartic acid residues at positions 443 and 498 respectively in the RNase H domain of HIV-1 RT were investigated by the defined substitution of these residues using site-directed mutagenesis. Two mutant HIV-1 RT enzymes were created using this technique, a D443A single mutant and a D443A/D498N double mutant. The substitution of Asp443 for alanine was specifically selected because the methyl group of the alanine residue, unlike the carboxylate group of aspartic acid, is unionisable and is unable to metal ions within the region of the RNase H active site. Hence, the alanine residue is unable to partake directly in general acid/base catalysis, or indirectly in electrophilic catalysis by the positioning of the divalent metal ion near the phosphoryl centre. The creation of the double mutant, D443A/D498N, was carried out in order to examine the possible structural relationship of the aspartic acid residues at these positions, as well as to clarify the findings of Mizrahi *et al.* (1990) with regard to the stabilising effect of the distal D443N mutation on the D498N substitution as had been observed in studies of the D443N/D498N mutant.

The construction and purification of the D443A and D443A/D498N mutants is described in the "Results" section. In the case of the D443A mutant, the final protein product of this process was shown to be indistinguishable from the wild-type HIV-1 RT enzyme as shown by preliminary RT activity analysis and SDS-PAGE gel electrophoresis (Figure 21). This finding showed that

the Asp443 in the wild-type enzyme is not essential for the structural integrity of RT since substitution of this residue for the alanine residue did not appear to adversely affect the expression or the gross structure of the enzyme. In addition the mutation would appear to show that any interaction that may occur in the wild-type enzyme between the carboxylate-bearing side chains of Asp443 and Asp498 is not vital to the maintenance of the enzyme's structure. The result also shows that the D443A mutation had no affect on cleavage by HIV-1 protease at the nearby Phe440/Tyr441 cleavage site which results in the formation of the p51 subunit of the p66/p51 heterodimer. This finding is consistent with the observation that the D443N mutant did not affect cleavage by HIV-1 protease at the p51/p15 boundary (Mizrahi *et al.*, 1990).

The D443A/D498N mutant enzyme showed no detectable RT activity and was not purifiable although restriction analysis of the final AR120/pDPTPRO4+pGALKRTE construct showed that the D443A/D498N mutant was indistinguishable from the wild-type and D443A mutant at the DNA level (Figure 14). The reason for the lack of any detectable expression of the D443A/D498N mutant RT enzyme is most likely due to the high stringency of the co-expression system. The high level of HIV-1 protease expression during induction, and the relatively promiscuous nature of HIV-1 protease cleavage, results in an environment that is highly selective against improperly folded proteins, a factor that makes this expression system particularly sensitive to structural abnormalities brought about by the substitutions made. The result

obtained with respect to the double mutant would suggest that the lack of recoverability of the enzyme is caused by a folding defect which renders the polypeptide unstable under the *in vivo* co-expression system. This folding defect must be a direct consequence of the substitutions made in this particular mutant and more specifically indicates that Asp498 plays an essential role in the formation of the correct tertiary structure for the entire HIV-1 RT protein. In addition, this finding, when considered in relation to the observed stability of the D443N/D498N RT mutant of Mizrahi *et al.* (1990) and the instability of this group's D498N mutant, implies that an interaction between the two asparagine residues, possibly through the formation of two C=O---H-N hydrogen bonds, is responsible for the stabilisation of this particular mutant. In light of the observed stability of the D443A mutant, it would therefore appear that the residues Asp443 and Asp498 of the wild-type enzyme are not required to directly interact with one another in order to maintain the structural stability of the enzyme and that the stability of the D443N/D498N mutant is not a true reflection of a direct structural relationship between Asp443 and Asp498 in the wild-type structure, but is rather a fortuitous surrogate stabilising interaction facilitated by their spacial proximity.

The ability of the D443A enzyme to carry out RNA-dependent DNA polymerisation was assessed in more detail by the determination of the K_m and K_{cat} values for polymerisation using a poly(rA).dT_{12-18} homopolymer substrate. The values obtained for these parameters showed that there was no significant difference

between the wild-type and D443A RT enzymes. This suggested that the D443A mutation within the RNase H region of HIV-1 RT had no influence on the polymerase activity of the enzyme and hence further confirmed the functional independence of the polymerase and RNase H activities of HIV-1 RT. This observation agrees with the observations of Dudding *et al.* (1991) who found that the mutations D443N and D443N/D498N had no effect on these kinetic parameters of polymerisation when using the same substrate. However, it should be noted that further investigation of RNA-dependent DNA polymerisation by these mutants revealed that there were significant differences in polymerisation when a heteropolymer RNA.DNA hybrid was used as a substrate (Dudding and Mizrahi, 1993). In this case, although the mutant and wild-type enzymes appeared to share the same pause sites during polymerisation, the product distribution of polymerisation was found to be significantly different when compared to that of the wild-type enzyme. Further investigation of the RNA-dependent DNA polymerase activity of the D443A mutant is therefore required, particularly with regard to its ability to synthesise a DNA strand using a heteropolymer RNA.DNA hybrid, before it can be said that the D443A mutation has no distal effect at all on the polymerase activity of the RT enzyme.

The investigation of the thermal stability of the D443A RT mutant revealed that the mutation had no significant effect on the mutant's thermal stability when compared to that of the wild-type enzyme. This finding can be expanded to conclude that the substitution of alanine for aspartic acid at position 443 results

in a tertiary RT protein structure that is as stable as that of the wild-type enzyme. The D443A mutation therefore has no apparent adverse effect whatsoever on the structure of the enzyme, unlike the D443A/D498N mutation which disrupted the tertiary structure to such an extent that recovery of this mutant from the co-expression system was not possible. Thermal stability studies carried out on the mutant RT enzymes D443N and D443N/D498N revealed that these enzymes were significantly less stable than the wild-type enzyme, a result that indicates a disruptive influence of these substitutions on the structural integrity of the RT enzyme (C.Nkabinde and V.Mizrahi, unpublished data). However, the greater structural instability of the D443N and D443N/D498N mutants, as observed in this case, was not sufficient to produce folding defects that prevented protein recovery and purification as was found for the D498N (Mizrahi *et al.*, 1990) and D443A/D498N mutants.

Examination of the RNase H activity of the D443A mutant revealed that the enzyme lacked any detectable RNase H activity, even when a 10 fold higher concentration of the enzyme was used in the assay. The result is probably a consequence of the inability of the alanine residue in the D443A mutant to bind a divalent metal ion at the 443 position, a role apparently associated with aspartic acid in the wild-type RNase H enzyme (Davies *et al.*, 1991). Divalent metal ions have been shown to be essential for RNase H activity (Oyama *et al.*, 1989). The result confirms the finding that Asp443 is vital to RNase H activity and it is suggested that its role in this activity is to position the

HIV-1 RT mutant enzyme.	Polymerase activity.	RNase H activity.
D443N[a]	+++	-
D443A	+++	-
D443E[b]	-	-
D498N[a]	-	-
D498E[b]	-	-
D443N/D498N[a]	+++	-
D443A/D498N	-	-
D443E/D498N[b]	-	-

Table 5: The mutations of Asp443 and Asp498 of HIV-1 RT carried out to date and the effects of these mutations on the polymerase and RNase H activities of the mutant RT enzymes. (a) Mizrahi *et al.*, 1990; (b) C.Nkabinde and V.Mizrahi, unpublished data.

divalent metal ion in proximity to the phosphoryl centre involved in hydrolysis at the RNase H catalytic site. The significance of this with regard to the mechanism of cleavage of the RNA moiety of an RNA.DNA hybrid by HIV-1 RT is discussed at the end of this section.

Broader conclusions regarding the RNase H active site of HIV-1 RT require a more detailed examination of the mutagenic and crystallographic data relating to the structure of this region. Table 5 is a summary of all the substitutions of Asp443 and Asp498 carried out to date. Figure 25 shows the Mn^{2+} binding site of the HIV-1 RT RNase H domain as determined by Davies *et al.* (1991). The figure shows that the two Mn^{2+} ions are tightly co-ordinated by four acidic residues, one ion probably being co-ordinated by Asp443 and Asp549 and the other by Asp498 and Glu478. It should also be noted that, whereas Asp443 lies within a ß-sheet, Asp498 is located within an α-helix.

A summary of the data in Table 5 shows three basic trends:

1) That the substitution of Asp443 for a residue that is isosteric (as in the case of asparagine) or sterically smaller (as in the case of alanine) results in complete loss of the RNase H activity, although polymerase activity is unimpaired on a homopolymer template and the resultant protein is purifiable.

2) That the substitution of Asp443 for a residue that is sterically larger but possesses a carboxylate group (as in the

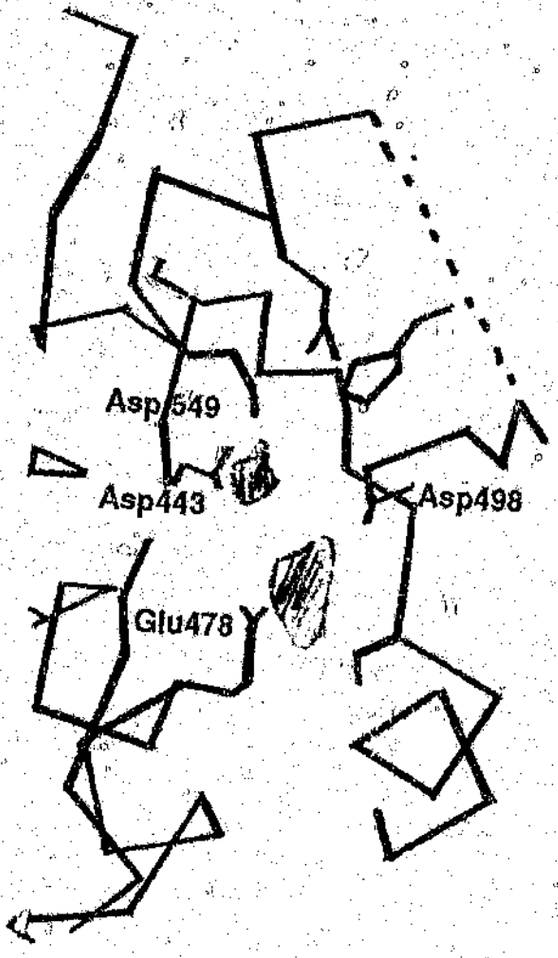


Figure 25: The Mg^{2+} binding site of the RNase H domain of HIV-1 RT. The positions of the catalytically important acidic residues Asp443, Glu478, Asp498 and Asp549 are indicated as well as those of the two bound Mg^{2+} ions. The dotted line represents the loop region containing His439 that was disordered in the crystal structure reported. (Adapted from Davies *et al.*, 1991.).

case of glutamic acid (D443E)) results in an unstable protein that is unpurifiable.

3) That any substitution of Asp498 results in an enzyme that is unstable and hence unpurifiable except in the case of the D443N/D498N double mutant. This, as discussed previously, may be the result of the stabilising influence, through hydrogen bonding, of the distal Asn443 residue on the otherwise unstable Asn498 structure.

The substitutions made at the Asp443 position reveal that the gross structure around this residue is flexible enough to allow the insertion of other relatively small side chains at this position without any great disruption of stability, a feature that may be due to the flexibility of the ß-sheet in which the 443 position is situated. However spacial constraints at this location are sufficient to bring about disruption when a larger side chain, as in the case of glutamic acid, is inserted. This is probably due to a steric clash with the side chain of the Asp549 residue, its normal partner in Mg^{2+} binding. In addition the results obtained for the thermal stabilities of the D443A and D443N mutants would appear to indicate that the small alanine residue is easily accommodated at this position and has a negligible effect on stability, whereas the larger asparagine residue has some destabilising effect, possibly by causing an alteration in the local electrostatic network of the protein. The loss of RNase H activity seen in these Asp443 mutants is probably due to the loss of the divalent metal ion binding capability at

this position. The Asp443-associated Mg^{2+} ion is therefore catalytically important, but it is not essential for the maintenance of a correctly folded structure.

In the case of Asp498 this residue would seem to play a very important role in maintaining the structure of the RT enzyme although it may also be involved in catalysis. Any substitution of this residue carried out in the co-expression system results in a protein that is unstable under the induction conditions. The exact nature of this phenomenon is remains unclear; however two features of the 498 position, or a combination thereof, may be of influence in this case. Firstly the Asp498 residue is located within a relatively rigid α-helix motif, a structure that imposes more severe spacial constraints upon the residue found at the 498 position. Secondly, it is possible that the second metal ion that appears to be coordinated by Asp498 and Glu478 is important to the structure of the protein and loss of the metal ion binding capacity of the residue at position 498 may result in a destabilisation of the enzyme. The possible role of the Asp498 residue in catalysis is still not known since no single mutation at this position has resulted in a purifiable enzyme and the D443N/D498N mutant was not informative in this respect because of the background RNase H phenotype of the D443N mutation.

As discussed in section 1.4.2, hydrolysis of the phosphodiester bond of RNA by retroviral RNase H results in the production of oligoribonucleotides bearing 3'-OH and 5'-phosphate termini. This finding suggests a catalytic mechanism that is

different from that of most other nucleases that leave free 5'-OH and 3'-phosphate termini. The available mutagenic and mechanistic data for the RNase H active site of HIV-1 RT supports a similar mechanism to that of the *E.coli* RNase HI activity, as would be expected given the structural similarity of these two proteins. However, Katayanagi *et al.* (1990;1992) and Davies *et al.* (1991) presented conflicting data with regard to the number of divalent metal ions bound at the RNase H active site. As discussed previously, Katayanagi *et al.* (1990;1992) showed that only one metal ion was co-ordinated at this position in *E.coli* RNase HI, whereas the data of Davies *et al.* (1991) suggested that two metal ions were co-ordinated at the HIV-1 RNase H active site. Both groups have offered explanations for this discrepancy, however the problem has still not been resolved.

Figure 26 shows a schematic diagram of a proposed mechanism for the action of the HIV-1 RT RNase H, similar to that proposed by Nakamura *et al.* (1992) for *E.coli* RNase HI, that requires only a single Mg^{2+} ion to be co-ordinated by the Asp443 residue. The diagram shows three acidic residues Asp443, Asp498 and Glu478 surrounding the phosphodiester bond of the RNA backbone. Cleavage of the P-O-3' bond is brought about by nucleophilic attack by a hydroxyl group after either Asp498 or Glu478 has acted as a proton acceptor from a water molecule (residue B). With the present body of knowledge, this role cannot be specifically attributed to either one of these residues. The nucleophilic attack on the P-O-3' bond is aided by the electrophilic nature of the Mg^{2+} ion which is coordinated by Asp443 and the remaining

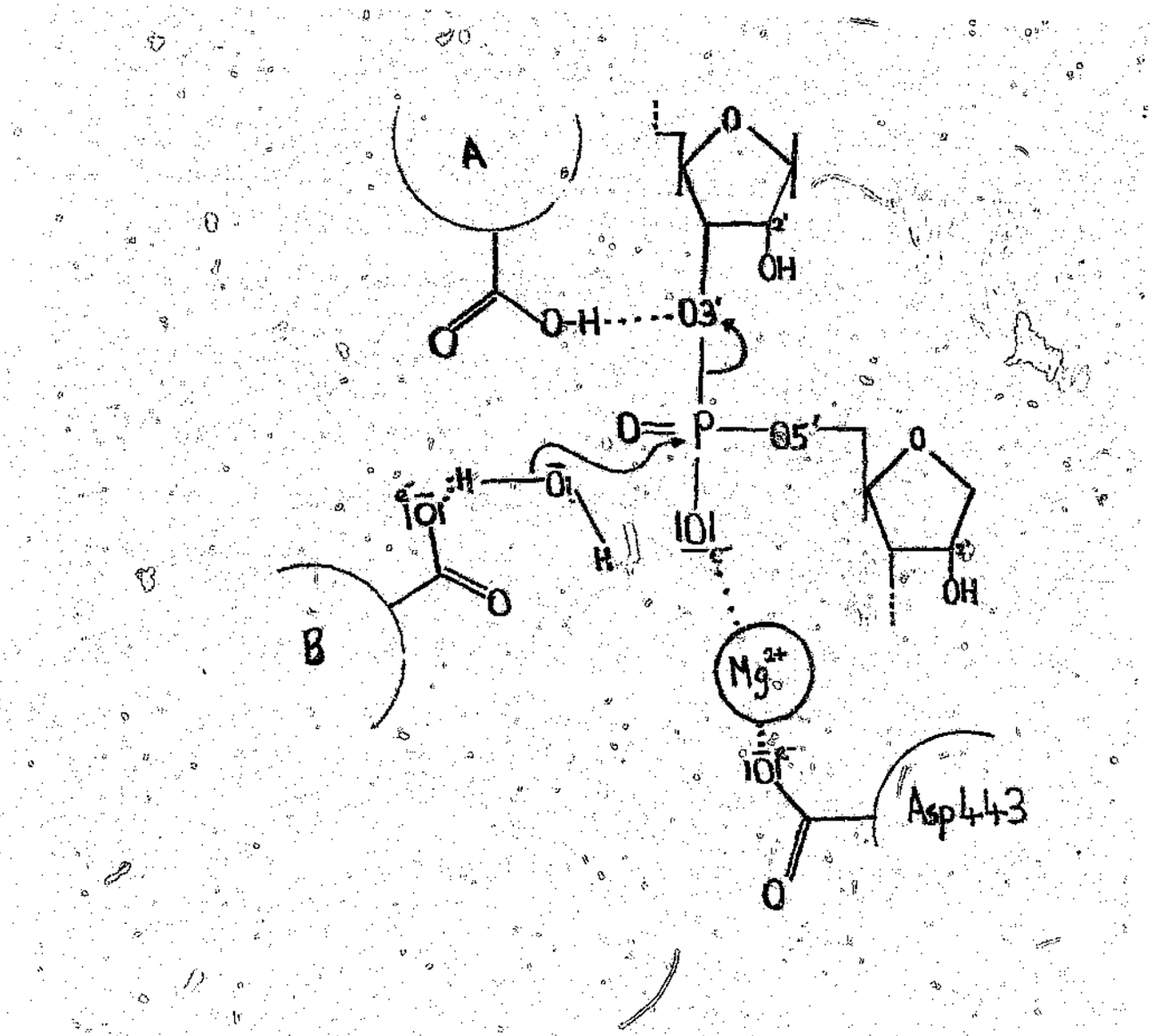


Figure 25: Schematic diagram for the proposed mechanism of the HIV-RT associated RNase H activity. See text for details.

carboxylate residue (residue A) which again can be either Asp498 or Glu478.

If the findings of Davies *et al.* (1991) are correct, then a mechanism of cleavage by HIV-1 RNase H that involves two divalent metal ions must be considered. The co-ordination of the second metal ion by Asp498 and Glu478 would mean that the carboxylate groups of these residues would not be able to participate in RNA cleavage in the same manner as that proposed for the single metal ion mechanism. Catalytic mechanisms involving two divalent metal ions have been proposed for a number of enzymes including the zinc metalloenzymes. Of particular relevance is the mechanism of phosphoryltransfer catalysed by the 3'-5' exonuclease activity of *E.coli* DNA polymerase I as proposed by Beese and Steitz (1991). They proposed that the one metal ion facilitates the formation of the attacking nucleophilic hydroxide ion, and the second metal ion serves to facilitate the leaving of the 3'-OH group as well as stabilising the arrangement of the phosphate group to allow attack by the hydroxyl ion. Application of this mechanism to that of HIV-1 RNase H may imply that the two metal ions that may be co-ordinated by Asp443, Glu478, Asp498 and Asp549 are acting in a similar manner.

Further investigation of Asp443 and Asp498 is still required, especially with respect to the clarification of the apparent structural and possible catalytic roles of the Asp498 residue. To this end, a number of substitutions of the 498 position are currently being investigated with the hope that a

stable enzyme will result. If successful this will provide some insight into the function of this essential residue in the RNase H activity of HIV-1 RT.

5) APPENDICES

APPENDIX A.

Ligation Procedure.

Ligation mixtures were set up as follows:

0.5 µl M13-ΔRV vector (± 25 ng) or *Bam*H1/*Xba*1 digested M13mp11.

6.0 µl electroeluted 348-bp *Asp*718 fragment or electroeluted *Bam*H1/*Xba*1 fragment.

1.0 µl 10 x ligation buffer (as supplied by manufacturer).

1.0 µl 0.1 M DTT.

1.0 µl 10 mM ATP (pH 7).

1.0 µl T4 DNA ligase (1 unit as defined by manufacturer).

The mixture was incubated at 4°C for 16-20 hours and then quenched by the addition of 1 µl of 0.5 M EDTA. The complete reaction mixture was then used to transform JM101.

APPENDIX B.

Luria Broth (LB).

8 g Tryptone

5 g Yeast extract

5 g NaCl

Make up to one litre with H_2O and autoclave.

To make Luria agar (LA) add 15 g of agar.

20 X SSC Buffer.

17.53 g NaCl

8.82 g Sodium citrate.

Add 80 ml of H_2O and dissolve.

Adjust pH to 7 with 10 M NaOH and make up volume to 100 ml.

Autoclave to sterilise.

50 x Denhardt's Solution.

1 g Ficoll

1 g polyvinylpyrolidine

1 g bovine serum albumin

Make up to 100 ml with H_2O

5 x RT Assay Buffer.

714 µl 1 M Tris.HCl (pH 7.9)

714 µl 1 M KCl

85.7 µl 1 M $MgCl_2$

486.3 µl H_2O

5 x TBE Buffer.

54 g Tris

27.5 g boric acid.

20 ml 0.5 M EDTA (pH 8.0)

Make up to 1l with H_2O

APPENDIX C.

10% SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE) for Protein Analysis.

Preparation of Gels.

The following stock solutions were used for SDS-PAGE.

4 x Stacking Gel Buffer.

0.5 M Tris.HCl (pH 6.8)

8 x Resolving Gel Buffer.

3.0 M Tris.HCl (pH 8.8)

30% Acrylamide Stock Solution.

29.2 g Acrylamide

0.8 g Bis-acrylamide

H_2O added to final volume of 100 ml.

10 x Reservoir Buffer Stock.

30.3 g Tris

144.0 g glycine

10.0 g SDS

A Bio-Rad BRL vertical V16-2 electrophoresis apparatus (Bio-Rad, Richmond, California, USA) with 20 x 30 cm glass plates and 0.8 cm spacers was used for all SDS-PAGE gels.

Resolving Gel.

6 ml 30% acrylamide stock

2.25 ml 8 x resolving gel buffer

0.18 ml 10 % SDS

0.9 ml fresh 1% ammonium persulphate

8.67 ml H_2O

10 µl TEMED

The mixture was mixed well and immediately poured between glass plates until it reached ± 2 cm below the top of the smaller plate. Water was gently layered over the top of the mixture and the gel allowed to polymerise for approximately 15 minutes. The water was then removed and the space between the plates dried.

Stacking Gel.

0.75 ml 30% acrylamide stock

1.25 ml 4 x stacking gel buffer

0.05 ml 10% SDS

0.25 ml 1% ammonium persulphate

2.7 ml H_2O

10 µl TEMED

The mixture was mixed well and immediately poured on top of the polymerised resolving gel to the top of the glass plate. A sample spacer comb was immediately inserted between the plates and the gel allowed to polymerise. The comb was then removed and the glass plates containing the gel was attached to the electrophoresis apparatus and 1 x reservoir buffer was poured in

the upper and lower buffer reservoirs such that the gel surface was in contact with the reservoir buffer.

Sample Preparation.

3 x Sample Loading Dye.

3.75 ml 0.5M Tris.HCl (pH 6.8)

0.6 g SDS

1.5 ml 2-mecaptoethanol

3.0 ml glycerol

0.24 ml 0.25% bromophenol blue

1.5 ml H_2O

10 µl of the sample loading dye was added to each 20 µl of protein samples, and the samples heated at 95°C for 3 minutes to denature protein. The samples were then loaded on to the gel and initially electrophoresed at 90 V for 60 minutes followed by 130 V until completion or overnight at 50 V.

Staining of Gels.

Following electrophoresis the gel was removed from between the two glass plates and stained at room temperature for 30 minutes using a solution of 0.5% Coomasse Brilliant Blue, 41.6% (v/v) methanol, and 16.7% (v/v) glacial acetic acid in H_2O. The gel was destained with 30% (v/v) methanol and 10% (v/v) glacial acetic acid and stored in 5% glycerol.

6.) REFERENCES.

Argos,P., (1988) *Nature* 226, 1209-1211

Arnold,E., Jacobo-Molina,A., Nanni,R.G., Williams,R.L., Xiaode,L., Ding,J., Clark,A.D.Jnr., Zhang,A., Ferris,A.L., Clark,P., Hizi,A., Hughes,S.H., (1992) *Nature* 357, 85-89

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

Baltimore,D., Smoler,D., (1971) *Proc. Natl. Acad. Sci. U.S.A* 68, 1507-1511

Barat,C., Lullien,V., Schatz,O., Keith,G., Nugeyre,M.T., Grüninger-Leitch,F., Barré-Sinoussi,F., LeGrice,S.F.J., Darlix,J.L., (1989) *EMBO J.* 8, 3279-3285

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

Becerra,S.P., Kumar,A., Lewis,M.S., Widen,S.G., Abbotts,J., Karawya,E.M., Hughes,S.H., Shiloach,J., Wilson,S.H., (1991) *Biochemistry* 30, 11707-11719

Beese,L.S., Steitz,T.A., (1991) *EMBO J.* 10, 25-33

Ben-Artzi,H., Zeelon,E., Gorecki,M., Panet,A., (1992A) *Proc. Natl. Acad. Sci. U.S.A* 89, 927-931

Ben-Artzi,H., Zeelon,E., LeGrice,S.F.J., Gorecki,M., Panet,A., (1992B) *Nucleic Acids Res.* 20, 5115-5118

Berg,J.M., (1986) *Science* 232, 485-487

Cann,A.J., Karn,J., (1989) *AIDS* 3 (suppl 1), S19-S34

Centers for Disease Control (CDC), (1981) *Morbid. Mortal. Wkly. Rep.* 30, 250-252

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

Clapham,P.R., Weber,J.N., Whitby,D., (1989) *Nature* 337, 368-370

Coffin,J.M., (1990) in *Virology* eds. Fields,B.N., Knipe,D.M., Channock,R.M., Melnick,J., Roizman,B., Slope,R., 2nd edition (Raven press, NY) 2, 1437-1450

Crouch,R.J., Dirksen,M.L., (1985) in *Nucleases* eds. Linn,S.M., Roberts,R.J., (Cold Spring Harbor Laboratory Press, Cold Sring Harbor, NY)

Davies,J.F.II, Hostomska,Z., Hostomsky,Z., Jordan,S.R., Mathews,D.A., (1991) *Nature* 252, 88-95

DeStefano,J.J., Buiser,R.G., Mallaber,L.M., Bambara,R.A., Fay,P.J., (1991A) *J. Biol. Chem.* 266, 24295-24301

DeStefano,J.J., Buiser,R.G., Mallaber,L.M., Bambara,R.A., Fay,P.J., (1991B) *J. Biol. Chem.* 266, 7423-7431

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

Dirani-Diab,R.E., Andreola,M-L., Nevinsky,G., Tharaud,D., Barr,P.J., Litvak,S., Tarrago-Litvak,L., (1992) *FEBS Letts.* 301, 23-28.

Donahue,P.R., Hoover,E.A., Beltz,G.A., Riedel,N., Hirsch,V.M., Overbaugh,J., Mullins,J.I., (1988) *J. Virol.* 62, 722-731

Doolittle,R.F., Feng,D-F., Johnson,M.S., McClure,M.A., (1989) *Q. Rev. Biol.* 64, 1-30

Dudding,L.R., Harington,A., Mizrahi,V., (1990) *Biochem. Biophys. Res. Commun.* 167, 244-250

Dudding,L.R., Nkabinde,N.C., Mizrahi,V., (1991) *Biochemistry* 30, 10498-10496

Dudding,L.R., Mizrahi,V., (1993) *Biochemistry* (in press)

Fennie,C., Lasky,L.A., (1989) *J.Virology* 63, 639-646

Fujinaga,K., Parsons,J.T., Beard,J.W., Beard,D., Green,M., (1970) *Proc. Natl. Acad. Sci. U.S.A* 67, 1432-1439

Fujiwara,T., Craigie,R., (1989) *Proc. Natl. Acad. Sci. U.S.A* 86, 3065-3069

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

Gallaher,W.R., (1987) *Cell* 50, 327-328

Gallo,R.C., Salahuddin,S.Z., Popovic,M., Shearer,G.M., Kaplan,M., (1984) *Science* 224, 500-503

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

Gilboa,E., Goff,S., Shields,A., Yoshimura,F., Mitra,S., Baltimore,D., (1979B) *Cell* 16, 863-874

Goff,S.P., Lobel,L.I., (1987) *Biochim. Biophys. Acta.* 907, 93-123

Goff,S.P., (1990) *J. Acquired Immune Defic. Syndr.* 3, 817-831

Gopalakrishnan,V., Peliska,J.A., Benkovic,S.J., (1992) *Proc. Natl. Acad. Sci. U.S.A* 89, 10763-10767

Gottlinger,H.G, Sodroski,J.G., Haseltine,W.A., (1989) *Proc. Natl. Acad. Sci. U.S.A* 86, 5781-5785

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

Gregerson,D.S., Albert,J., Reid,T.W., (1980) *Biochemistry* 19, 301-306

Gross,L., (1951) *Proc. Soc. Exp. Biol. Med.* 78, 27-32

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

Harada,F., Peters,G.G., Dahlberg,J.E., (1975) *J. Biol. Chem.* 254, 10979-10985

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

Hizi,A., Hughes,S.H., Shaharabany,M., (1990) *Virology* 175, 326-337

Hostomsky,Z., Hostomska,Z., Hudson,G.O., Moomaw,E.W., Nodes,B.R., (1991) *Proc. Nat. Acad. Sci. U.S.A.* 88, 1148-1152

Hostomsky,Z., Hudson,G.O., Rahmati,S., Hostomska,Z., (1992) *Nucleic Acids Res.* 20, 5819-5824

Hsu,T.W., Sabran,J.L., Mark,G.E., Guntaka,R.V., Taylor,J.M., (1978) *J. Virol* 28, 810-818

Hu,W., Temin,H.M., (1990) *Science* 250, 1227-1230

Huber,H.E., Tabor,S., Richardson,C.C. (1987) *J. Biol. Chem.* 262, 16224-16232

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

Inouye,S., Hsu,M-Y., Angle,S., Inouye,M., (1989) *Cell* 56, 709-717

Jacobo-Molina,A., Clark,A.D.Jnr., Williams,R.L., Nanni,R.G., Clark,P., Ferris,A.L., Hughes,S.H., Arnold,E., (1991A) *Proc. Natl. Acad. Sci.* 88, 10895-10899

Jacobo-Molina,A., Arnold,E., (1991B) *Biochemistry* 30, 6351-6361

Ji,J., Loeb,L.A., (1992) *Biochemistry* 31, 954-958

Johnson,M.S., Mclure,M.A., Feng,D-F., Gray,J., Doolittle,R.F., (1986) *Proc. Nat. Acad. Sci. U.S.A.* 83, 7648-7652

Jones,K.A., Kadonaga,J.T., Luciw,P.A., Tjian,R., (1986) *Science* 232, 755-759

Kacian,D.L., Watson,K.F., Burny,A., Spiegelman,S., (1971) *Biochem. Biophys. Acta.* 246, 365-383.

Kalyanaraman,V.S., Sarngadharan,M.G., Robert-Guroff,M., Miyoshi,I., Blayney,D., (1982) *Science* 218, 571-573

Kamer,G., Argos,P., (1984) *Nucleic Acids Res.* 12, 7269-7282

Kanaya,S., Kohara,A., Miura,Y., Sekiguchi,A., Iwai,S., Inoue,H., Ohtsuka,E., Ikehara,M., (1990) *J. Biol. Chem.* 265, 4615-4621

Katayanagi,K., Miyagawa,M., Ishikawa,M., Matsushima,M., Kanaya,S., Ikehara,M., Matsuzaki,T., Morikawa,K., (1990) *Nature* 347, 306-309

Katayanagi,K., Miyagawa,M., Matsushima,M., Ishikawa,M., Kanaya,S., Nakamura,H., Ikehara,M., Matsuzaki,T., Morikawa,K., (1992) *J. Mol. Biol.* 223, 1029-1052

Keller,W., Crouch,R., (1972) *Proc. Natl. Acad. Sci. U.S.A* 69, 3360-3363

Khan,R., Giedroc,D.P., (1992) *J. Biol. Chem.* 267, 6689-6695

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

Kowalski,M., Potz,J., Basiripour,L., Dorfman,T., Wei,C.G., (1987) *Science* 237, 1351-1355

Krug,M.S., Berger,S.L., (1989) *Proc. Natl. Acad. Sci. U.S.A* 86, 3539-3543

Kunkel,T.A., Roberts,J.D., Zakour,R.A., (1987) *Methods Enzymol.* 74, 5463-5467

Lampson,B.C., Sun,J., Hsu,M-Y., Vallejo-Ramerez,J., Inouye,S., Inouye,M., (1989) *Science* 243, 1033-1038

Larder,B.A., Purifoy,D.J.M., Powell,K.L., Darby,G., (1987) *Nature* 327, 716-717

Larder,B.A., Kemp,S.D., Purifoy,D.J., (1989A) *Proc. Natl. Acad. Sci. U.S.A* 86, 4803-4807

Larder,B.A., Kemp,S.D., (1989B) *Science* 246, 1155-1158

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

Le Grice,S.F.J., Naas,T., Wohlgensinger,B., Schatz,O., (1991) *EMBO J.* 10, 3905-3911

Leis,J., Berkower,I., Hurwitz,J., (1973) *Proc. Natl. Acad. Sci. U.S.A* 70, 466-470

Lever,A., Gottlinger,H.G., Heseltine,W.A., Sodroski,J.G., (1989) *J.Virol* 63, 4085-4087

Levy,J.A., Hoffman,A.D., Kramer,S.M., Landis,J.A., Shimabukuro,J.M., (1984) *Science* 225, 840-842

Lori,F., Scovassi,A.I., Zella,D., Achilli,G., Cattaneo,E., Casoli,C., Betazzoni,U., (1988) *AIDS Res. Hum. Retrovir.* 4, 393-398

Merluzzi,V.J., Labadia,M., Grozinger,K., Skoog,M., Wu,J.C., Shih,C-K., Eckner,K., Hattox,S., Adams,J., Rosenthal,A.S., Faanes,R., Eckner,R.J., Koup,R.A., Sullivan,J.L., (1990) *Science* 250, 1411-1414

Mervis,R.J., Ahmad,N., Lillehoj,E.P., Paum,E.J., Salazai,F.H.R., (1988) *J.Virol* 62, 3993-4002

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

Mizrahi,V., (1989B) *Biochemistry* 28, 9088-9094

Mizrahi,V., Usdin,M.T., Harington,A., Dudding,L.R., (1990) *Nucleic Acids Res.* 18, 5359-5363

Moelling,K., Bolognesi,D.P., Bauer,H., Büsen,W., Plassmann,H.W., Hausen,P., (1971) *Nature New Biology* 234, 240-243

Nabel,G., Baltimore,D., (1987) *Nature* 326, 711-713

Nagashunmugam,T., Velpandi,A., Goldsmith,C.S., Zaki,S.R., Kalyanaraman,V.S., Srinivasan,A., (1992) *Proc. Natl. Acad. Sci. U.S.A* 89, 4114-4118

Nakamura,H., Oda,Y., Iwai,S., Inoue,H., Ohtsuka,E., Kanaya,S., Kimura,S., Katsuda,C., Katayanagi,K., Morikawa,K., Miyashiro,H., Ikehara,M., (1991) *Proc. Natl. Acad. Sci. U.S.A* 88, 11535-11539

Ollis,D.L., Brick,P., Hamlin,R., Xuong,N.G., Steitz,T.A., (1985) *Nature* 313, 762-767

Omer,C.A., Faras,A.J., (1982) *Cell* 30, 797-805

Omer,C.A., Resnick,R., Faras,A.J., (1984) *J. Virol* 50, 465-470

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

Paganiban,A.T., Fiore,D., (1988) *Science* 241, 1064-1069

Panet,A., Haseltine,W.A., Baltimore,D., Peters,G., Harada,F., Dahlberg,J.E., (1975) *Proc. Natl. Acad. Sci. U.S.A* 72, 2535-2539

Peliska,J.A, Benkovic,S.J., (1992) *Science* 258, 1112-1118

Poiesz,B.J., Ruscetti,F.W., Gazdar,A.F., Bunn,P.A., Minna,J.D., (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77, 7415-7419

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

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

Prats,A.C., Sarih,L., Gabus,C., Litvak,S., Keith,G., Darlix,J.L., (1988) *EMBO J.* 7, 1777-1783

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,J.S.R., Pearson,M.L., Lautenberger,J.A., Papas,T.S., Ghrayeb,J., Chang,N.T., Gallo,R.C., Wong-Staal,F., (1985) *Nature* 313, 227-284

Rattray,A.J., Champoux,J.J, (1987) *J. Virol.* 61, 2843-2851

Rattray,A.J., Champoux,J.J, (1989) *J. Mol. Biol.* 208, 445-456

Reardon,J.E., Miller,W.H., (1990) *J. Biol. Chem.* 265, 20302-20307

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

Rho,H.M., Poiesz,B., Roscetti,F.W., Gallo,R.C., (1981) *Virology* 112, 355-360

Rosen,C.A., (1991) *Trends in Genetics* 7, 9-14

Roth,M.J., Schwatzberg,P.L., Goff,S.P., (1989) *Cell* 58, 47-54

Sallafranque-Andreola,M.L., Robert,D., Barr,P.J., Fournier,M., Litvak,S., Sarih-Cottin,L., Tarrago-Litvak,L., (1989) *Eur. J. Biochem* 184, 367-274

Sambrook,J., Fritsch,E.F., Maniatis,T., (1989A)in *Molecular Cloning, a laboratory manual*, 2nd edition (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), volume 1, p 4.29-4.30.

Sambrook,J., Fritsch,E.F., Maniatis,T., (1989B)in *Molecular Cloning, a laboratory manual*, 2nd edition (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), volume 1, p 6.9-6.13.

Sambrook,J., Fritsch,E.F., Maniatis,T., (1989C)in *Molecular Cloning, a laboratory manual*, 2nd edition (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), volume 1, p 1.25-1.28.

Sambrook,J., Fritsch,E.F., Maniatis,T., (1989D)in *Molecular Cloning, a laboratory manual*, 2nd edition (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), volume 1, p 1.82-1.84.

Sanchez-Pescador,R., Power,M.D., Barr,P.J., Steimer,K.S., Stempien,M.M., Brown-Shimer,S.L., Gee,W.W., Renard,A., Randolph,A., Levy,J.A., Dina,D., Luciw,P.A., (1985) *Science* 226, 484-492

Sarngadharan,M.G., Popovic,M., Bruch,L., Schupbach,J., Gallo,R.C., (1984) *Science* 224 503-505

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

Schatz,O., Cromme,F.V., Naas,T., Lindemann,D., Mous,J., LeGrice,S.F.J, (1990A) in *Gene Regulation and AIDS* ed. Papas,T.S., (Portfolio Publishing Company, Houston, USA) 293-303

Schatz,O., Mous,J., LeGrice,S.F.J, (1990B) *EMBO J.* 9, 1171-1176

Schupbach,J., Popovic,M., Gilden,R., Gonda,M.A., Sarngadharan,M.G., (1984) *Science* 224, 503-505

Shank,P.R., Hughes,S.H., Kung,H-J., Majors,J.E., Quintrell,N., Guntaka,R.V., Bishop,J.M., Varmus,H.E., (1978) *Cell* 15, 1386-1396

Shinnick,T.M., Lerner,R.A., Sutcliffe,J.G., (1981) *Nature* 293, 543-548

Skalka,A.M., (1989) *Cell* 56, 911-913

Skinner,J.A., Eperon,I.C., (1986) *Nucleic Acids Res.* 14, 6945-6964

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

South,T.L., Kim,B., Summers,M.F., (1989) *J. Am. Chem. Soc.* 111, 395-396

Starnes,M.C., Gao,N., Ting,R.Y.C., Cheng,Y-C., (1987) *J. Biol. Chem.* 263, 5132-5134

Starnes,M.C., Cheng,Y-C., (1989) *J. Biol. Chem.* 264, 7073-7077

Stein,B.S., Gowda,S.D., Lifson,J.D., Penhallow,R.C., Bensch,K.G., (1987) *Cell* 49, 659-668

Stein,H., Hausen,P., (1969) *Science* 166, 393-395

Tanese,N., Goff,S.P., (1988) *Proc. Natl. Acad. Sci. U.S.A* 85, 1777-1781

Taylor,J., Shameen,L., (1987) *J. Cell Sci.* 7 (suppl), 189-195

Tisdale,M., Schulze,T., Larder,B.A., Moelling,K., (1991) *J. Gen. Virol.* 72, 59-66

Vaishnav,Y.N., Wong-Staal,F., (1991) *Annu. Rev. Biochem.* 60, 577-630

Veronese,F.D., Copeland,T.D., Oroszlan,S., Gallo,R.C., Saranagadharan,M.G., (1988) *J.Virol* 62, 795-801

Wain-Hobson,S., Sonigo,P., Danos,O., Cole,S., Alizon,M., (1985) *Cell* 40, 9-17

Wakefield,J.K., Jablonski,S.A., Morrow,C.D. (1992) *J. Virol.* 66, 6806-6812

Watson,K.F., Schendel,P.L., Rosok,M.J., Ramsey,L.R., (1979) *Biochemistry* 18, 3210-3219

Weiss,R., Teich,N., Varmus.H.E., and Coffin,J., (1982) *RNA Tumor Viruses*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Wintersberger,U., (1990) *Pharmacology and Therapeutics* 48, 259-280

Wong,S.F., Wahl,A.F., Yuan,P-M., Arai,N., Pearson,B.E., Arai,K., Korn,D., Hunkapiller,M.V., Wang,T.S-F., (1988) *EMBO J.* 7, 37-47

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

Yu,H., Goodman,M.F., (1992) *J. Biol. Chem.* 267, 10888-10896

Author: Brooksbank Richard Leslie.
Name of thesis: The role of the conserved ASP443 and ASP498 residues in the polymerase and RNase H activities of HIV-1 reverse transcriptase.

PUBLISHER:
University of the Witwatersrand, Johannesburg
©2015

LEGALNOTICES:

Copyright Notice: All materials on the University of the Witwatersrand, Johannesburg Library website are protected by South African copyright law and may not be distributed, transmitted, displayed or otherwise published in any format, without the prior written permission of the copyright owner.

Disclaimer and Terms of Use: Provided that you maintain all copyright and other notices contained therein, you may download material (one machine readable copy and one print copy per page)for your personal and/or educational non-commercial use only.

The University of the Witwatersrand, Johannesburg, is not responsible for any errors or omissions and excludes any and all liability for any errors in or omissions from the information on the Library website.