Application of *in vitro* transcribed and translated HIV-1 reverse transcriptase and integrase in protein-protein interaction studies

by

Line Wergeland

A thesis submitted in partial fulfillment of the requirements
for the degree of *Candidata Scientiarum*

Department of Molecular Biology
University of Bergen
December 2000

Contents

Acknowledgements

The work for this thesis has been carried out in Lab. 1 at the Department of Molecular Biology, University of Bergen from January 1999 to December 2000. My supervisor has been dr. philos Dag E. Helland. I am very grateful for his help and guidance during the work with this thesis. I am also deeply indebted to dr. philos Bjørn-Ivar Haukanes for taking the time to read my manuscript.

I also want to thank everybody at Lab. 1 for giving me all the technical help and advice I needed at all stages of creating this thesis. They also contributed to a very enjoyable and inspiring work environment.

Many thanks also go to Annicke Stranda, Elisabeth Siem, Karoline Bragstad and Åsne Jul-Larsen for numerous discussions and lunchbreaks.

Finally I want to thank Torbjørn Jacobsen, my family and friends for all the support and encouragement they have given me.

Bergen, November 2000

Line Wergeland

Abbreviations

AIDS	Acquired Immunodeficiency Syndrome
CA	HIV-1 Capsid Protein
CAT	Chloramphenicol Acetyl Transferase
CITE	Cap-Independent Translation Enhancer
ECL	Enhanced Chemiluminiscense
ELISA	Enzyme Linked Immunosorbent Assay
EMCV	Encephalomyocarditis Virus
ExPASy	Expert Protein Analysis System
gp	glycoprotein
HIV-1	Human Immunodeficiency Virus type 1
HRP	Horseradish Peroxidase
IN	HIV-1 Integrase
LB	Luria Bertani
LMP agarose	Low Meltingpoint Agarose
LTR	Long Treminal Repeats
MA	HIV-1 Matrix Protein
MLV	Murine Leukaemia Virus
NC	HIV-1 Nucleocapsid Protein
NLS	Nuclear Localisation Signal
NMR	Nuclear Magnetic Resonance
OD	Optical Density
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Primer Binding Site
PCR	Polymerase Chain Reaction
PIC	HIV-1 Preintegration Complex
POD	Horseradish Peroxidase
PPT	Polypurine Tract
PR	HIV-1 Protease
R-region	Repeat-region
RT	HIV-1 Reverse Transcriptase
SDS	Sodium Dodecyl Sulphate
SIB	Swiss Institute of Bioinformatics
SPR	Surface Plasmon Resonance
SU	Surface Subunit
TM	Transmembrane Subunit
U3-region	Unique 3’ region
U5-region	Unique 5’ region

Summary

Replication of the human immunodeficiency virus (HIV) necessitates the two viral enzymes reverse transcriptase (RT) and integrase (IN). These two proteins are found to associate in a large nucleoprotein complex called the preintegration complex (PIC) together with viral nucleic acids and other viral and host proteins. Since the two proteins are localised together in the host cell, a direct interaction between them can exist, although it has been difficult to demonstrate such interaction. In this study, RT and IN were produced by *in vitro* transcription and translation and the proteins were used to study possible RT-IN interactions by ELISA.

The RT and IN coding regions were amplified by PCR from a full length HIV-1 HxB2 clone and ligated in-frame into the *in vitro* translation vector pCITE-5b(+). The two plasmids generated, pCITE-RT and pCITE-IN were constructed and verified by sequencing and restriction enzyme analysis.

A coupled *in vitro* transcription and translation kit based on rabbit reticulocyte lysate was utilised together with the two plasmids in order to obtain RT and IN. The proteins were produced unmarked since marked amino acids possibly could obscure an eventual interaction. Verification of the protein products was done by Western blotting having specific monoclonal antibodies. The proteins produced were not purified from the reticulocyte lysate but used directly in an ELISA-format to try to detect any interactions between RT and IN.

With the ELISA-format, no interactions could be demonstrated. The signal was the same for all dilutions of the coating protein, as well as for the negative control. The most likely explanation for this is that there are some proteins in the reticulocyte lysate that interacts with both RT and IN.
If *in vitro* translated proteins are going to be used to investigate protein-protein interactions by this ELISA method they need to be purified first. An other possibility is to cotranslate the two proteins *in vitro*. It is not sure that an eventual direct interaction between RT and IN can be detected by ELISA. Other components of the PIC may need to be present for a direct interaction to occur.

1 Introduction

The human immunodeficiency virus type 1 (HIV-1) is the causative agent of acquired immunodeficiency syndrome (AIDS). HIV-1 is a retrovirus of the lentivirus subfamily.

1.1 Genomic organisation of HIV-1

The genome of HIV-1 consists of two identical single-stranded RNA molecules which is converted to double stranded DNA by the viral enzyme reverse transcriptase (RT). Long terminal repeats (LTRs) are flanking the viral genes on the double-stranded DNA. The HIV-1 genome has three major coding regions; *gag*, *pol* and *env*. The *gag* gene encodes the virion capsid proteins, the *pol* gene encodes the viral enzymes needed for replication (protease (PR), integrase (IN) and RT) and the *env* gene encodes the envelope glycoprotein. In addition to the *gag*, *pol* and *env* coding regions, the HIV-1 encodes several other proteins. This includes Vif and Vpr that are part of the viral particle, Tat and Rev that regulates viral gene expression and Vpu and Nef that interacts with the cellular machinery to promote virus propagation [1]. An overview of the viral genome and virion structure is given in Figure 1.

The virion has a spherical shape and is about 110 nm in diameter. It consists of a lipid bilayer envelope that surrounds the cone-shaped nucleocapsid. Computer simulations and scanning electron microscopy has shown that the virion is icosahedral [2, 3], although the overall shape of the virion is spherical. The nucleocapsid consists of the viral RNA molecules surrounded by the Gag proteins; MA which is located between the lipid envelope and the nucleocapsid, CA which forms the capsid shell and NC which binds to the RNA. A $tRNA^{Lys3}$ primer is located inside the nucleocapsid, together with the viral enzymes PR, RT and IN. The viral accessory proteins Vif and Vpr are also present in the viral particle, but their exact location is still unknown.

The viral envelope contains a number of knobs that shows triangular symmetry [4]. A knob is thought to consist of four heterodimers of the Env gp, each heterodimer is composed of a surface subunit (SU, gp120) and a transmembrane subunit (TM, gp41).

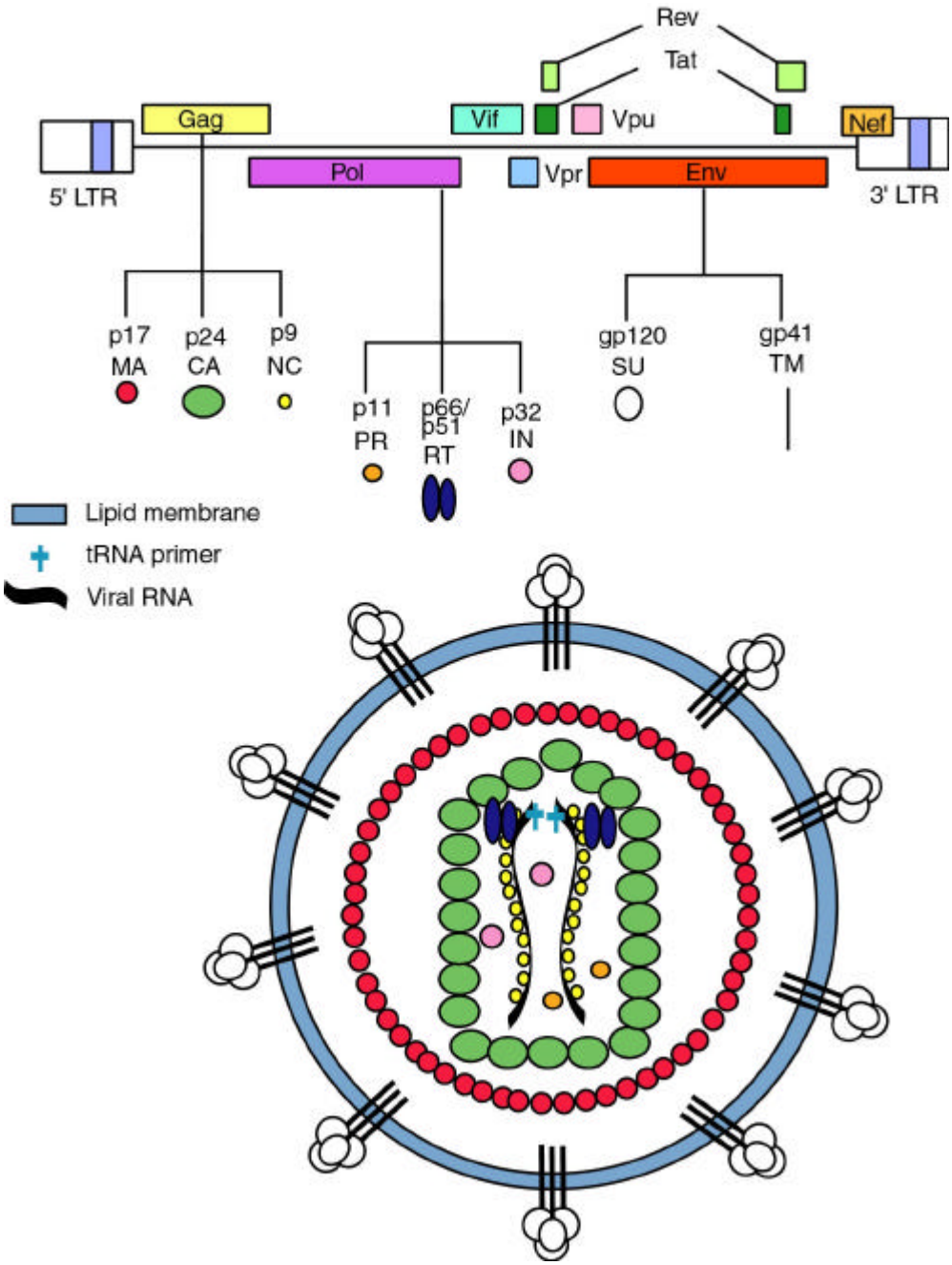


Figure 1. The HIV-1 virion structure. The precursor polyproteins Gag, Pol, and Env are enzymatically cleaved to yield the mature virion proteins. The Gag and Pol polyproteins are cleaved by the viral protease (PR) to produce the following proteins; the matrix protein (MA), the major capsid protein (CA), the nucleocapsid protein (NC), protease (PR), reverse transcriptase (RT) and integrase (IN). The Env polyprotein is cleaved by a cellular protease to give the surface subunit (SU) gp 120 and the transmembrane subunit (TM) gp 41 of the surface Env glycoprotein. The HIV-1 genome also has open reading frames for the regulatory proteins Tat and Rev, and the accessory proteins Vif, Vpr, Vpu and Nef.

Biochemical and immunochemical analyses have also shown that the HIV membrane contains several cellular proteins acquired during budding [5]. Their role in viral replication has not been resolved.

1.2 The HIV-1 lifecycle

In order to simplify the rather complex lifecycle of HIV-1, a schematic overview is presented in Figure 2. A more detailed description of each step is given in the chapters below.

1.2.1 Binding and entry

The HIV-1 virus has the CD4 molecule as its major receptor. Entry of HIV-1 into target cells starts with an interaction between gp120, CD4 and a seven-transmembrane G protein coupled chemokine coreceptor [6]. Binding of gp 120 to CD4 induces conformational changes and increases the exposure of gp41 epitopes allowing the host cellular membrane to fuse with the lipid-envelope coat of the virion [7-9]. Following fusion, the virion core is released into the cytoplasm of the host cell (Figure 2).

At this stage, the viral RNA is a part of a large nucleoprotein complex in which the reverse transcription occurs. This complex contains at least RT, IN and a portion of the MA protein [10-13] and is also known as the preintegration complex (PIC). See 1.3 for details.

1.2.2 Reverse transcription

Reverse transcription of the viral RNA starts from a $tRNA^{Lys3}$ primer bound to a primer binding site (PBS) at the viral RNA. The enzyme reverse transcriptase (RT) copies the 5' R-U5 (repeat-unique 5') region of the RNA into "minus strand strong stop" DNA and the RNase H activity of RT degrades the RNA. This degradation induces the first strand transfer, where the newly synthesised DNA hybridises to the complementary 3' R (repeat) region of either the same RNA strand or on the other copy of viral DNA. Then the minus strand DNA synthesis continues through the viral genome. Again RNase H degrades the RNA in the RNA-DNA hybrid, but leaves a 16-base polypurine tract (PPT) 5' to the U3 (unique 3') region. This PPT is used to prime the synthesis of the plus strand DNA. When the synthesis reaches the $tRNA^{Lys3}$ primer, it stops and the primer is removed by RNase H. Then the second strand transfer reaction occurs and the newly synthesised PBS at the plus strand hybridises with the

complementary PBS on the minus strand. Finally RT completes the synthesis of both the strands.

1.2.3 Nuclear transport and integration

The double stranded DNA remains attached to the nucleoprotein complex, from this stage called the preintegration complex (PIC). A detailed description of this complex will be given later (see 1.3.1). This complex is transported to the cell nucleus where the integration occurs. For details see 1.3.2.

After the DNA has reached the nucleus, integrase (IN) is capable of integrating the viral DNA into the host genome. IN first removes two nucleotides from the 3' end of both DNA strands, producing recessed 3' ends [14]. Then it cuts the host DNA, creating 5' overhangs. The integration reaction is completed by cellular DNA repair enzymes, which ligate the ends and fill the gaps. The viral DNA is now integrated in the host DNA as a provirus. Cellular signals convert the cell from the latent stage to a virus producing stage. The mechanisms for this conversion are unknown.

1.2.4 Expression of viral proteins

The integrated provirus is flanked by tandem LTRs (see Figure 1). The 5' LTR contains an enhancer-promotor, and the 3' LTR contains a polyadenylation site, similar to a eucaryotic gene. Transcription of the provirus is done by the cellular polymerase II, resulting in a primary transcript. The primary transcript is mostly spliced, but some full-length transcripts occur at this early phase. There are two important functions of the full-length primary transcript; it serves as the viral genomic RNA that is packaged into virons and it serves as mRNA for the translation of all the viral proteins.

The HIV-1 promotor is highly regulated by both cellular and viral factors. In a model for regulation of viral gene expression, the expression is separated into two phases. There is the early phase, where multiply spliced mRNAs are synthesised. These mRNAs encodes the non virion proteins Tat, Rev and Nef. This phase is followed by the late phase where an accumulation of unspliced and singly spliced mRNAs leads to production of the proteins that are incorporated into virions [15, 16].

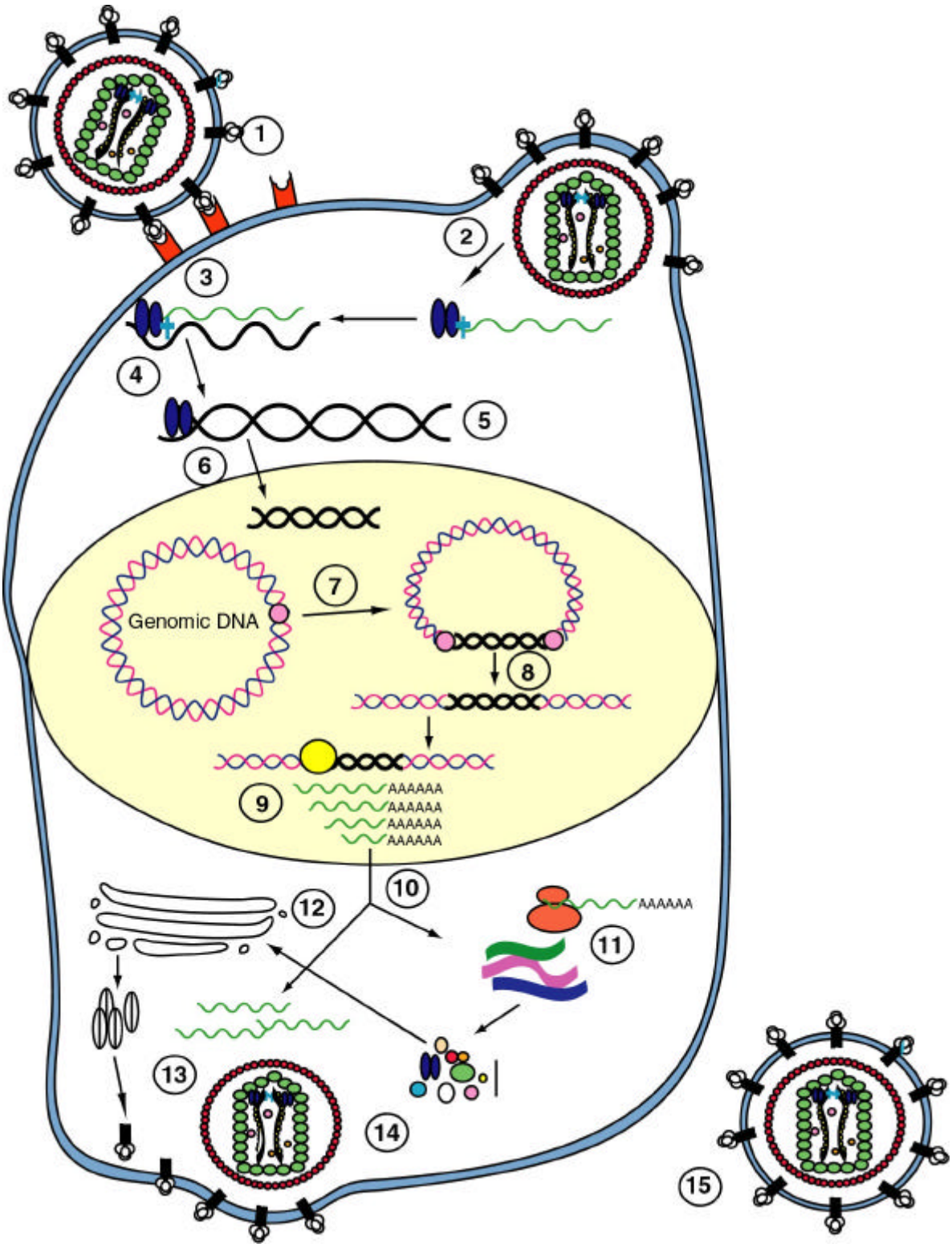


Figure 2. The viral lifecycle. (1) The virus binds to the CD4 receptor on the cell surface by an interaction between gp120 and CD4. (2) The host cellular membrane fuses with the virion membrane and the virion core is released into the cytoplasm. (3) DNA is produced from viral RNA by reverse transcription. (4) RNaseH degrades the RNA. (5) The second strand of DNA is produced by RT. (6) The preintegration complex migrates to the nucleus. (7) The cellular DNA is cut, and the viral DNA integrated into the host genome by integrase. (8) The viral DNA is latent in the host genome. (9) Host enzymes transcribe the cellular genes. (10) The RNA is transported out of the nucleus and processed. (11) Viral polyproteins are synthesised. (12) Some viral proteins are glycosylated. (13) The viral RNA and viral proteins are assembled into the virion. (14) The virion buds from the plasma membrane. (15) Maturation of the virion.

The viral protein Tat activates the transcription of the HIV-1 genes and leads to high levels of spliced transcripts early after productive infection i.e. the early phase. The Rev protein regulates the balance of spliced to unspliced and partially spliced mRNAs. Activation by Rev leads to the late phase of viral expression.

1.2.5 Virion assembly, budding and release

Prior to virion assembly the major structural proteins accumulate inside the plasma membrane. Then the Gag polyprotein oligomerises [17] and the Gag protein is transported to the plasma membrane. It is the matrix (MA) domain of the Gag protein that is responsible for this transport [18]. The Gag polyprotein complex starts forming the inner framework of the virion and two copies of the viral RNA assemble around a motif in the NC domain of the Gag polyprotein [19]. The encapsidation process is completed when the $tRNA^{Lys3}$ primer associates with the Pol domain of the Gag-Pol polyprotein and is included in the virion [20]. Maturation of the virion occurs as a coordinated process by which the viral protease cleaves the Gag-Pol polyprotein in an ordered fashion [21]. Following maturation, the Gag and Gag-Pol polyproteins begin to encapsidate the viral RNA and the plasma membrane forms a lipid bilayer around the virion. The gp160 is glycosylated and oligomerised in the endoplasmatic reticulum [22, 23] and transported to the Golgi apparatus. Here the glycoprotein is processed into gp120 (SU) and gp41 (TM) [24]. Then they associate noncovalently and are transported to the cell surface [25]. Finally the inner core shell condenses and is linked with the lipid membrane containing the viral glycoproteins.

1.3 The preintegration complex (PIC)

Following cell infection by HIV-1, the viral proteins remain associated with the viral nucleic acids in a high molecular weight nucleoprotein complex (preintgration complex, PIC) [12, 26]. After reverse transcription the viral DNA must be transported to the cell nucleus before it can be integrated into the host genome. This transport of PIC to the nucleus is achieved by means of a process requiring ATP, but independent of cell division. [27]. The composition of PIC is relatively well established, but its means of nuclear transport is still under extensive investigation.

1.3.1 The composition of PIC

Several different approaches have been used to purify the HIV-1 PIC. The results of these differing approaches are fairly uniform. The virion protein most stably associated with PIC is IN [12]. RT, MA and Vpr can also be detected under a number of conditions [10, 11, 13, 28-30]. Under very limited conditions NC, PR and low amounts of CA have also been detected [10, 13, 28]. The HIV-1 PIC does not contain high amounts of capsid proteins, a fact that distinguishes it from the murine leukaemia virus (MLV) PIC [10-13]. A number of cellular proteins have been shown to associate with PIC. These include importin-α [29, 31]; HMG I(Y) [32], which is important for PIC activity *in vitro*; and histones [10].

1.3.2 Nuclear transport of PIC

Transport of the HIV-1 PIC to the nucleus during interphase is a unique feature that distinguishes HIV-1 and the lentiviruses from other retroviruses. The lentiviruses are not dependent on the dissolution of the nuclear envelope during mitosis to get access to the cellular genome [33, 34]. Nuclear import of PIC is in addition believed to be critical for HIV-1 replication in non-dividing and slowly dividing cells [30, 35-37]. There are three components of the HIV-1 PIC that are believed to be important for the nuclear import; IN, Vpr and MA.

IN has been shown to have a role in nuclear import of PIC [29]. IN has a bipartite nuclear localisation signal (NLS) that is recognised by the importin/karyopherin pathway. A combination of mutations in this NLS region of IN, in Vpr and in the NLS of MA totally eliminated the nuclear import of the HIV-1 PIC in P4 cells, but with virus that was only defective in Vpr and the MA NLS, the import was just partially reduced. This was taken as proof for the importance of IN in the process of nuclear import.

Vpr is another component of the HIV-1 PIC that is implicated in nuclear import [30, 35]. Vpr bind to karyopherin α, but to a different site than the NLS binding site. This binding increases the affinity of the karyopherin α and NLS interaction [38]. Vpr has also been shown to interact with some nucleoporins [39] and to regulate binding of PIC to the nuclear pore complex [40, 41]. It is not known whether Vpr can function as an independent HIV-1 nuclear import factor.

MA is the third and most important member of PIC when it comes to nuclear import. An NLS in the N-terminal region of MA was first identified [37, 42] and believed to determine the nuclear translocation activity of HIV-1 PIC. This hypothesis was later

disrupted, when mutations in this NLS had little effect on HIV-1 replication in non-dividing cells [43, 44]. It has also been shown that the nuclear import function is maintained when MA has a large N-terminal truncation [45]. These findings questioned the role of MA in the nuclear translocation process. Recently a new NLS in MA has been discovered [46], called NLS-2. The NLS-2 is located in the C-terminal region of MA. When both the N-terminal NLS and NLS-2 in MA were inactivated, the virus was rendered unable to replicate, despite the presence of both Vpr and IN. This indicates that MA is necessary for the efficient nuclear import of the HIV-1 PIC *in vivo*.

1.4 Structure and function of reverse transcriptase

Reverse transcriptase (RT) is a heterodimer consisting of one subunit with molecular weight of 66 kDa (p66), and one subunit of 51 kDa (p51). The protein is processed from the Gag-Pol polyprotein by PR during virion assembly [47]. The processing of Gag-Pol occurs in two steps; first the p66 is cleaved from the polyprotein, and a homodimer is formed. Second, one of the subunits in the p66 homodimer is cleaved near the C-terminus and the heterodimer is formed.

Several structures of RT are reported, all found by X-ray crystallography. They share the same structural features, and differ only in the molecules used together with RT in the crystal. A schematic model of RT is given in Figure 3.

Kohlstaedt,*et al* describes an enzyme-inhibitor complex [48]. They divide the polymerase regions of p66 and p51 into four subdomains denoted “fingers”, “palm”, “thumb” and “connection” (see Figure 3). The subdomains are arranged differently in each subunit, thus the heterodimer is asymmetric. This is remarkable, since p51 is formed from a C-terminal deletion of p66. The major contacts between p51 and p66 occur in the connection domains. The thumb of p51 contacts the RNase H domain of p66. This interaction may be required for RNase H activity [49]. In addition the palm of p66 is in contact with the fingers of p51. In the p51 subunit, the connection domain is rotated such that it occupies the palm and buries the active site residues. This is consistent with the fact that mutations in the active site of p51 have little effect on RT activity [50]

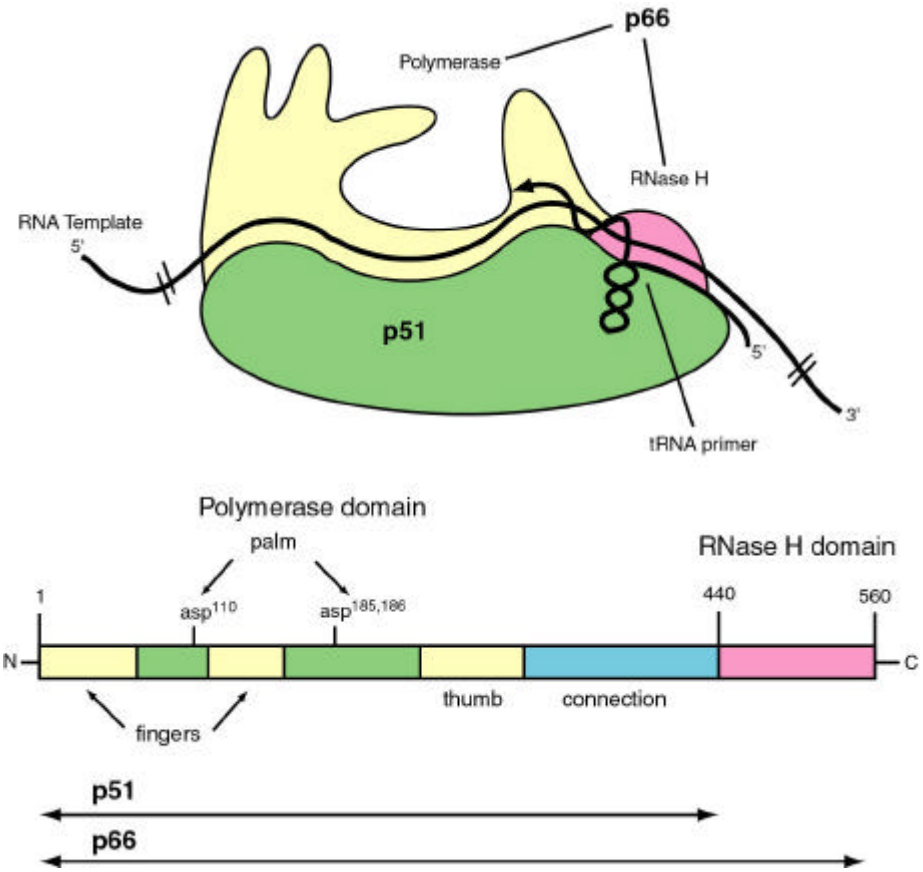


Figure 3. An overview of the RT domains. After PR processes RT from the Gag-Pol precursor to yield a p66 homodimer, a part of the C-terminus is cleaved from one of the subunits to produce the p66/p51 heterodimer. The structures of the subdomains of p66 have been compared with a right hand and the subdomains are subsequently designated as palm, thumb and fingers. A connection domain joins the polymerase domain to the RNase H domain. Asp^{110} and $Asp^{185,186}$ are residues involved in catalysis. The viral RNA and the tRNA primer are positioned in the palm.

An other crystal structure determined for RT [51], describes a DNA-template-primer complex bound in the cleft region. The cleft is flanked by a putative nucleoside triphosphate binding site and the RNase H active site. This is consistent with models that suggest that the polymerase active site and the RNase H active site act coordinately during the synthesis of minus strand DNA [52]. A higher resolution structure has also been reported by the same group [53]. This shows the putative catalytic residues Asp^{110}-Asp^{185}-Asp^{186} positioned near the primer strand 3' hydroxyl. In this structure , the bound template-primer shows both A- and B-form regions of DNA separated by a bend. Contacts between the template strand and the palm and thumb were also noted.

RT is an enzyme with three important enzymatic activities:

-RNA-dependent DNA polymerase

-DNA-dependent DNA polymerase

-Ribonuclease H (RNase H)

These activities are used to copy the single-stranded RNA-genome of the virus into a double-stranded DNA molecule that can be integrated into the host genome. The polymerase activity of RT lies within the N-terminal portion of p66 and the RNase H activity resides in the C-terminal part of the same subunit.

The catalytic site of RT lies in a cleft in the palm of the p66 subunit (see Figure 3). It contains the sequence Tyr^{183}-Met^{184}-Asp^{185}-Asp^{186} (YMDD) in addition to Asp^{110}. Two α-helices in the thumb of the p66 subunit together with the palm subdomain act as a clamp to position the template-primer relative to the polymerase active site. The 3' hydroxyl of the primer terminus is held close to the catalytically essential Asp^{110}-Asp^{185}-Asp^{186} residues, thus in position for nucleophilic attack on incoming nucleoside triphosphates. The connection subdomain of p66 interacts with both the template and p51.

1.5 Structure and function of integrase

Integrase (IN) is an enzyme of 31 kDa, produced from the C-terminal part of Pol after the processing of the Gag-Pol polyprotein by the HIV-1 protease (PR).

IN comprises three independently folding domains; an N-terminal domain, a catalytic core domain and a C-terminal domain (see Figure 4). The structures of the separate domains are known, but the manner in which they interact in the holoenzyme is still not known. This is because IN has not yet been crystallised, due to its poor solubility.

The first 50 amino acids define the N-terminal domain. A conserved HHCC motif, similar to certain zinc-binding motifs, is the most significant region of homology in this family of proteins [54]. The isolated N-terminal of IN has actually been shown to bind zinc at a 1:1 molar ratio [55, 56]. The zinc-binding is required for folding of the isolated N-terminal domain and also enhances multimerisation of the full-length IN [56, 57]. The exact function of the N-terminal domain is not known. It has been suggested that this domain is involved in DNA-binding [58], but deletion of the N-terminal domain does not affect the DNA-binding activity of IN [59].

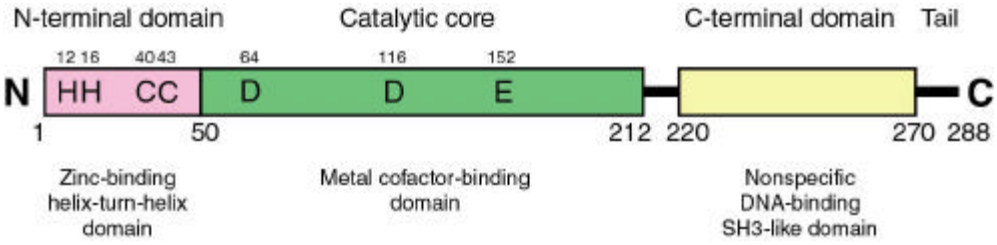


Figure 4. Model of HIV-1 IN. This model shows the three IN folding domains with known structure. The region called "tail" refers to the last 18 amino acids at the C-terminus. The structure of this region is not known. The conserved and catalytically important residues are indicated with corresponding residue numbers. This figure is adapted from Asante-Appiah and Skalka [60].

The structure of the N-terminal domain of IN has been solved by NMR [61]. It is composed of four α-helices with a hydrophobic core stabilising the lower region of the monomer. The upper part of the monomer is stabilised by the coordination of zinc to the HHCC motif (see Figure 5a). The N-terminal domain is a dimer with a hydrophobic dimer interface. This dimer resembles several helix-turn-helix DNA binding domains. It is nevertheless not obvious that the N-terminal domain of IN is binding DNA. The second helix of the IN N-terminal domain is part of the dimer interface as opposed to the corresponding DNA binding helix in helix-turn-helix DNA domains.

The catalytic core domain is a well conserved domain among retroviral integrases and contains a DD(35)E motif. The residues in this motif; Asp^{64}, Asp^{116} and Glu^{152}, are key residues in the active site [62, 63]. It has been proposed that these residues coordinate divalent metal ions critical for catalysis [62]. Mutations of any of these residues diminish or abolish the catalytic activity of IN [62, 64].

The structure of the catalytic core domain has been determined by X-ray crystallography [65]. As a dimer it is spherical in shape, with each monomer forming one hemisphere. Each monomer comprises a central five-stranded ß-sheet and six helices (see Figure 5b). The conserved acidic residues (Asp^{64}, Asp^{116} and Glu^{152}) are located in close proximity in the structure. But unexpectedly, the two catalytic sites are located on opposite sides of the spherical dimer, in a distance higher then the sites of insertion on the target DNA.

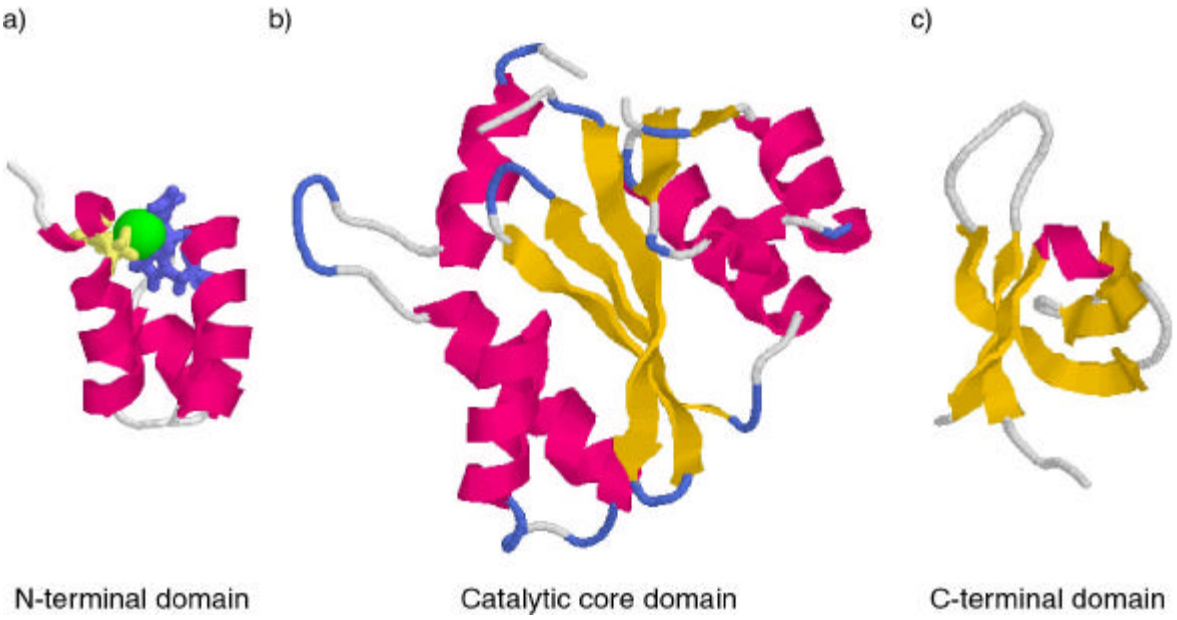


Figure 5. Structure of the different IN domains. a) Structure of a monomer of the N-terminal zinc-binding domain of IN. This structure is determined by NMR [61]. The zinc atom is showed in green and the coordinating His and Cys residues in blue and yellow respectively. b) Structure of a monomer of the catalytic core domain of IN. The structure of this domain is determined by X-ray crystallography [65]. c) Structure of a monomer of the C-terminal domain of IN. This structure is determined by NMR [66]. The three molecules are visualised using Raswin 2.6 and the files 1wjc, 1itg and 1qmc from Brookhaven National Laboratory Protein Data Bank.

The C-terminal domain is less conserved then the other domains. This region possesses non-specific DNA binding activity similar to that of the full-length integrase [67-71]. It has therefore been suggested that this domain is responsible for the binding of IN to the host genome.

The structure of the C-terminal domain was determined by using NMR [66, 72]. The structure is dimeric, with each monomer composed of a five-stranded ß-barrel (see Figure 5c). Stacking of ß-strands 2, 3 and 4 of each monomer in an antiparallell configuration forms the dimer. The dimer has a large saddle-shaped cleft that contains a number of positively charged residues. This cleft has the proper shape and size to accommodate double-stranded DNA.

IN mediates covalent linkage of linear double-stranded viral DNA to the host cell genome, thus possesses both DNA cleavage and joining activities. IN also mediates the reverse reaction i.e. disintegration *in vitro* [73]. This observation is consistent with the idea that the mechanism of integration involves a concerted mechanism and not a protein-DNA intermediate. Studies *in vitro* with deletion and site-specific point mutations in IN show that Asp^{64}, Asp^{116} and Glu^{152} are essential for cleavage, joining and disintegration [74]. This finding, together with the three-dimensional model for the catalytic domain of IN, indicates that this viral enzyme carries out all three reactions within a single active site.

1.6 The RT – IN interaction, and its study

Both RT and IN are known to be present in the HIV-1 PIC. In this complex, all the components need to be kept in close proximity. It is therefore very likely that direct interactions between RT and IN occur. An important question in this regard is where IN is located during the reverse transcription. Recent studies have shown that IN is essential for the initiation of reverse transcription in infected cells [75]. This strengthens the theory that there is a direct physical interaction between RT and IN. Several other observations also suggest that RT and IN may form a heterodimeric complex: The two proteins are known to coexist as a complex in some retroviruses [76, 77]. Further, the C-terminal domain of RT (RNase H) and the central core domain of IN are structurally similar [65, 78]. Finally, IN and RT can be coimmunoprecipitated with antibodies to either protein in murine leukaemia virus [79].

There are a number of different techniques developed to study protein-protein interactions. A description of all of them would be too extensive for this thesis. Therefor a selection of techniques believed to be applicable to the problem of detecting RT-IN interactions has been done. These techniques are described below.

1.6.1 Coimmunoprecipitation

Coimmunoprecipitation is a classical method of detecting protein-protein interactions and has been used in literally thousands of experiments. It exists a variety of different approaches that shares the common feature of using a precipitating antibody to test for physical interactions between proteins. In a typical experiment cells are lysed and a whole-cell extract is prepared under nondenaturing conditions. Following, the protein is precipitated form the lysate with a solid-phase affinity matrix (i.e. protein G-Sepharose) and the precipitate is tested for the presence of a second specifically associated protein. This method can be used for either native or epitope-tagged proteins for which antibodies are available. It is also possible to use recombinant proteins that have been engineered to bind with high affinity to a molecule that can be coupled to a solid-phase matrix. The presence of an associated protein is detected by separating the precipitated proteins by SDS-PAGE and Western blotting with an antibody recognising the putative associated protein. It is very important to add controls that test the specificity of interaction.

1.6.2 The yeast two-hybrid system

The yeast two-hybrid system is a yeast-based method to detect protein-protein interactions *in vivo* [80, 81]. It is based on the fact that many transcriptional activators consist of two modular domains; a DNA-binding domain and an activation domain. Normally these domains reside in the same protein, but it has been shown that separated domains can be able to assemble a functional transcriptional activator when brought into close proximity. The most used two-hybrid system is based on a GAL4 binding domain. The GAL4 protein is a transcriptional activator required for galactose utilisation in *Saccharomyces cerevisiae*. The N-terminal part of the protein is DNA-binding, and the C-terminal part activates transcription. Chimeric proteins can be made between the GAL4 domains and the proteins under investigation, using standard molecular biology cloning techniques. Introduction of the two plasmids encoding the chimeric proteins into a yeast strain with a reporter gene under control of the GAL4 activator allows for the detection of weak and transient interactions.

1.6.3 Physical chemical methods

BIAtechnology from Biacore AB is a physical-chemical approach for detecting protein-protein interactions. BIAtechnology enables detection of interactions between molecules, in real time, without the use of labels. BIAtechnology relies on the phenomenon of surface plasmon resonance (SPR) which occurs when surface plasmon waves are excited at a metal/liquid interface. Light is directed at, and reflected from, the side of the surface not in contact with sample, and SPR causes a reduction in the reflected light intensity at a specific combination of angle and wavelength. Molecular binding events cause changes in the refractive index at the surface layer, which are detected as changes in the SPR signal. In general, the refractive index change for a given change of mass concentration at the surface layer, is practically the same for all proteins and peptides, and is similar for glycoproteins, lipids and nucleic acids.

1.6.4 Enzyme Linked Immunosorbent Assays (ELISA)

ELISA was first described in 1971 as a method for detecting soluble antigens and antibodies. The assay method has been very successful owing to its sensitivity, speed and reproducibility of the assays. In addition to detecting soluble antigens and antibodies, ELISA can be used to detect protein-protein interactions. There are at least two different formats for detecting such

interactions. The first format involves coating with one of the proteins under investigation, blocking, adding the second protein, following an antibody for the second protein and addition of a secondary antibody conjugated to an enzyme. In case of an interaction, consequent addition of substrate will lead to development of colour. In the second format, the coating is done with an antibody for the first protein, following addition of the protein. The rest of the procedure is the same as in the first format. These two formats require small amount of protein, and are relatively easy to perform.

1.7 Aims of study

The HIV-1 proteins RT and IN are known to act together in the PIC. Several studies have been performed, trying to establish an interaction between the two proteins. By employing the yeast two-hybrid system, no such interaction could be detected [82]. Recently, however, an interaction has been detected with an in vitro binding assay. [75].

These are only two of several different ways to study protein-protein interactions. The sources of protein for these studies can also vary. The aims of this study are to produce HIV-1 RT and IN by *in vitro* transcription and translation and to see if the resulting proteins can be used for protein-protein interaction studies. Because the ELISA method requires small amounts of protein, this particular method was chosen.

2 Materials and methods

2.1 Enzymes

Name	Concentration	Supplier
*Bam*HI	10 U/µl	Promega
*Eco*RI	12 U/µl	Promega
Rnase	1 mg/ml	Boeringer Mannheim
Taq polymerase	5 U/µl	Promega
*Sac*I	10 U/µl	Promega
T4 DNA Ligase	3 U/µl	Promega

2.2 Plasmids

pCITE-5b(+) (3.7 kb) from Novagene (Figure 6).

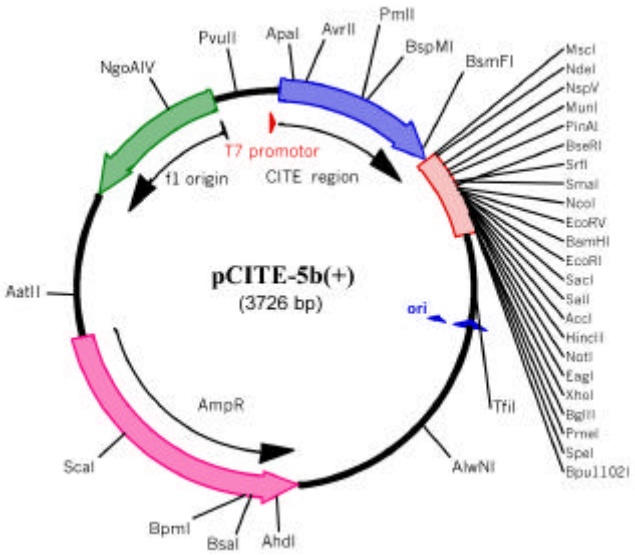


Figure 6. The plasmid pCITE-5b(+) from Novagene.

pSVC21 containing the entire HIV-1 genome from the HxB2 clone (gift from William A. Haseltine).

2.3 Antibodies

Name	Supplier
Anti-RT antibodies: 8C4, 11B7	Szilvay,*et al* [83]
Anti-IN antibodies: 8E5, 8G4	Nilsen,*et al* [84]
Biotin conjugated anti-mouse	Dako A/S
Horse radish peroxidase conjugated streptavidin	Amersham
Horse radish peroxidase conjugated anti-mouse	Amersham

2.4 Bacterial strains

E. coli Xl1Blue (recA1, endA1, gyrA96, thi, hsdR17 (rk[+], mk[+]), supE44, relA1, {lambda}[-], lac, [F', proAB, lacIqC{Delta}M15, tn10 (tet[R])])

2.5 Growth media

Bacto-Agar, peptone, tryptone and yeast extract are from Difco.

The ampicillin is from Sigma.

LB-medium (Luria-Bertani):

10 g Tryptone
5 g Yeast extract
10 g NaCl
ddH_2O were added up to 1 litre, and the solution was autoclaved.

Agar plates were made by adding 20 g/litre Bacto-agar before autoclaving the solution.

To make ampicillin plates, 50 µg/ml of ampicillin was added after autoclaving.

SOC-medium:

2% Tryptone
0.5 % Yeast extract
10 mM NaCl
2.5 mM KCl
10 mM $MgCl_2$
10 mM $MgSO_4$ } Autoclaved
20 mM Glucose

The $MgCl_2$, the $MgSO_4$ and the glucose are added from sterile 1 M stock solutions immediately before use.

2.6 Solutions and buffers

2.6.1 Produced solutions and buffers

Name	Description	Comment
0.5 M Tris-HCl, pH 6.8	6 g Tris is dissolved in 60 ml dH_2O pH is adjusted to 6.8 with HCl Water is added to 100 ml	
1 % Gelatine	10 g gelatine Dissolved in 1 litre 1 X PBS	
1 X Blotting buffer	7 parts of dH_2O 2 parts of Methanol 1 part of 10 X Blotting buffer	
1.5 M Tris-HCl, pH 8.8	27.23 g Tris is dissolved in 80 ml dH_2O pH is adjusted to 8.8 with HCl Water is added to 150 ml	
10 X Blotting buffer	144 g Glycine 30.3 g Tris Water is added to 1 litre	The solution is autoclaved
10 X PBS	80 g NaCl 2 g KCl 14.4 g Na_2HPO_4 2.4 g KH_2PO_4	
10 X Running buffer	60 g Tris base 288 g Glycine 20 g SDS Water is added to 2 litres	
10 X TBS	300 ml NaCl 200 ml Tris-HCl, pH 7.5 500 ml dH_2O	
30 % acrylamide, bis	87.6 g acrylamide 2.4 g bisacrylamide Water is added to 300 ml	

Name	Description	Comment
3M Sodiumacetate, pH 4.8	40.81 g sodium acetate 80 ml H_2O pH is adjusted to 4.8 with glacial acetic acid Water is added to 100 ml	
Blocking solution	1 g Ficoll 1 g Polyvinylpyrrolidon 10.3 g Boric acid 1.1 g NaOH 8.81 g NaCl 0.4 g NaN_3 0.5 g Gelatine 1 ml Nonidet P-40 (NP-40) 1 L dH_2O 15 g BSA is put on top	The solution is then put on a magnetic stirrer until most of the BSA is dissolved, and then filtrated.
GTE-solution	50 mM glucose 20 mM Tris-HCL pH 8 10 mM EDTA	The solution is autoclaved and stored at 4°C
KAc-solution	29.5 ml glacial acetic acid KOH pellets until a pH of 4.8 is reached H_2O to 100 ml	Stored at room temperature
NaOH-solution	0.2 N NaOH 1 % SDS	Freshly prepared
Sample buffer	200 mM Tris-HCL, pH 6.8 4 % SDS 20 % glycerol Pyronin B (saturated solution in ethanol) H_2O to 100 ml	
TE-buffer	10mMTris-HCL pH 8 1 mM EDTA pH 8	The solution is autoclaved

2.6.2 Purchased solutions and buffers

Name	Supplier
10 X Ligation buffer	Promega
10 X Multicore buffer	Promega
10 X PCR buffer	Promega
6 X Loading buffer	Promega
Big Dye	PE Biosystems
dATP	Promega
dCTP	Promega
dGTP	Promega
dTTP	Promega

2.7 Molecular weight standards

Name	Supplier
pGEM DNA	Promega
λ/*Hind*III DNA	Promega
Prestained Protein Marker Broad Range	New England Biolabs
HIV-1 IN (23 mg/ml)	Provided and purified by Professor Dag E. Helland
HIV-1 RT-His tag (0.75 mg/ml, 75% pure)	Provided and purified by Professor Dag E. Helland

2.8 Chemicals

Name	Formula	Supplier
2-Mercaptoethanol	C_2H_6OS	Merck
Acrylamide		Merck
Agarose NA		Pharmacia Biotech.
Ammonium persulphate		Biorad
Bisacrylamide		Biorad
Bovine Serum Albumine, (BSA)		Sigma
D(+)-Glucose-Monohydrate	$C_6H_{12}O_{6\cdot H2O}$	Merck
di-Sodium hydrogen phosphate dihydrate	Na_2HPO_4	Merck
Ethanol		Arcus produkter
Ethylenediaminetetraacetic acid, (EDTA)	$C_{10}H_{14}N_2O_8Na_{2\cdot 2H2O}$	Sigma
Ethylphenyl polyethylene glycol, (Nonidet P-40)		United States Biochemical corp.
Ficoll 400		Sigma
Gelatine		Difco
Glycerol	$C_3H_8O_6$	Merck
Glycine	H_2NCH_2COOH	Merck
Hydrochloric acid	HCl	Merck
Low-melting-point agarose		BRL
Magnesium chloride hexahydrate	$MgCl_{2\cdot 6H2O}$	Merck
Magnesium sulphate	$MgSO_4$	Merck
Methanol	CH_4O	Prolabo
N,N,N',N'-Tetramethylendiamin, (TEMED)	$C_6H_{16}N_2$	Merck
Polyoxyethylene sorbitan monolaurate, (Tween-20)		Sigma chemical co.
Polyvinylpyrrolidon		Merck
Potassium chloride	KCl	Merck
Potassium dihydrogen phosphate	KH_2PO_4	Merck
Potassium hydroxide pellets	KOH	Merck
Sodium azide	NaN_3	Riedel-De Haën ag.
Sodium chloride	NaCl	Merck
Sodium dodecyl sulphate, (SDS)	$CH_3(CH_2)_{11}SO_4Na$	United States Biochemical corp.
Sodium hydroxide pellets	NaOH	Merck
Tris	$NH_2C(CH_2OH)_3$	United States Biochemical corp.

2.9 Kits

Name	Supplier
TNT® T7 Quick Coupled Transcription / Translation system	Promega
Quiagen megaprep	Quiagen

2.10 Electrical equipment

Description	Name	Supplier
Centrifuge	Sorwall RC5B Refrigerated Superspeed Centrifuge	Sorwall
Developer	Fuji X-ray film developer RGII	Fuji
Electroporator	Gene Pulsar	Biorad
ELISA-plate reader	Multiscan MS	Labsystems
Eppendorf centrifuge	Centrifuge A 14	Jouan
Incubator	B5028	Heraeus
Incubator with shaking	HT	Infors AG
Incubator with shaking, eppendorf	Thermomixer	eppendorf
PCR-machine	2400 thermocycler	Perkin Elmer
Power supply	Model 200/2.0	Biorad
Sequencer	ABI PRISM 377 DNA Sequencer	Applied Biosystems
Spectrophotometer	UV-1201 uv-vis spectrophotometer	Shimadzu
Vacuum centrifuge	SpeedVac SVC100	Savant
Vortexer	Model K-550-GE	Bender & Hobein AG

2.11 Miscellaneous

Description	Supplier	Comment
Acrylamide gel equipment	Biorad	Mini-PROTEAN 3
Agarose gel equipment	Owl Separation Systems Inc.	Model B1
Autoradiography film	Amersham pharmacia biotech	Hyperfilm
Biotin-Lysyl-tRNA	Promega	Transcend™ tRNA
Blotting equipment	Biorad	Mini Trans-Blot Cell
ECL Western blotting detection reagents	Amersham Life Science	
Electroporation cuvettes	Biorad	Gene Pulser ® Cuvette
Filterpaper	Whatman	Chromatography paper 3MM Chp
Immunoplates	Nunc	F96 MaxiSorp Nunc-Immuno Plates
Nitrocellulose membrane	Schleicher & Schuell	Protran

2.12 Bacteriological methods

2.12.1 Preparation of electro-competent *E.coli* cells

One colony of *E.coli* Xl1Blue was inoculated in 5 ml of LB-medium and left to grow over night (37ºC, moderate shaking). 2.5 ml of this culture was inoculated in 500 ml LB-medium and left to grow at 37°C and with moderate shaking to an OD_{600} of approximately 0.5. The cells were chilled on ice, transferred to four pre-chilled GSA centrifuge tubes and centrifuged at 4200 rpm for 20 minutes at 4°C. The supernatants were poured off, and the pellets dissolved in 1.25 ml ice-cold water before adding 125 ml of water to each tube and centrifuging again as before. The supernatants were poured off again and the pellets dissolved in the remaining liquid. Another 125 ml of cold water was added, and the liquid was centrifuged under the same conditions. After the supernatants were poured off, the pellets were dissolved and 10 ml of ice-cold 10% glycerol was added. Another centrifugation was performed at 4200 rpm for 10 minutes at 4°C. The supernatants were poured off and the pellets dissolved in the remaining glycerol. The cells were then aliquoted and stored at –70°C.

2.12.2 Transformation of cells

The cells were transformed by electroporation. The electroporator (Biorad Gene Pulsar) was adjusted to deliver a pulse of 2.5 kV. An appropriate amount of plasmid DNA was mixed with 100 µl of electro-competent cells. The mixture was transferred to a pre-chilled electroporation cuvette and the cuvette was placed in the apparatus. A pulse was applied, following addition of 1 ml of SOC-medium. Then the mixture was transferred to a sterile eppendorf-tube and incubated for 45 minutes at 37°C and moderate shaking. 100 µl of the culture was plated out on a LB-plate containing ampicillin. The rest of the bacteria was spun down, and redissolved in 100 µl SOC-medium. This was also plated out on a LB-plate containing ampicillin. The plates were incubated at 37°C over night.

2.12.3 Isolation of plasmids from bacteria

Alkaline lysis minipreparation

For small-scale preparations of plasmid DNA (for detection and characterisation), the alkaline lysis method was used. One single bacterial colony was scraped from a LB-plate and inoculated into 5 ml of LB-medium containing the appropriate antibiotic. The culture was grown to saturation over night, at 37°C and moderate shaking. 1.5 ml of the culture was spun down for 20 seconds at 13000 rpm and the supernatant removed. The pellet was resuspended in 100 µl GTE-solution and left to stand for 5 minutes at room temperature. 200 µl of NaOH-solution was added, the solution was gently mixed by flicking the tube and the tube was placed on ice for 5 minutes. After this, 150 µl of KAc-solution was added, the mixture was vortexed and placed on ice for 5 minutes. The solution was then centrifuged for 3 minutes at 13000 rpm to pellet cell debris and chromosomal DNA. The supernatant was transferred to a fresh tube, mixed with 0.8 ml 96 % ethanol and left in room temperature for 2 minutes to precipitate nucleic acids. Centrifuging for 1 minute at 13000 rpm pelleted these nucleic acids. The supernatant was removed, and the pellet washed with 70 % ethanol. The pellet was dried in a vacuum centrifuge before resuspending in 30 µl TE-buffer.

The plasmids were characterised by restriction enzyme analysis (1 to 5 µl of plasmid solution was used), run on an agarose gel containing ethidium bromide and visualised under UV-light. Following, the gel was photographed and the picture kept for further analysis.

Quiagen megapreparation

To obtain a high yield of very pure plasmid, the Quiagen megaprep protocol was used. The protocol was followed as the manufacturer recommends.
The concentration of the plasmid solution was determined using agarose gel-electrophoresis and UV spectrophotometry.

The following formula was used to calculate the double stranded DNA concentration:

DNA concentration = A_{260} · 50 µg/ml · Dilution

2.13 PCR amplification of the RT and IN coding regions

2.13.1 The PCR reaction

The coding regions of RT and IN with different restriction enzyme cutting sites were amplified by PCR. The following primers were used:

Name	Sequence	Length	RE site
LWRT5	5' CC GGG GAT CCC ATT AGC CCT ATT GA 3'	25 bp	*Bam*H1
LWRT3	5' G GGA ATT CCG TAG TAC TTT CCT GAT TC 3'	27 bp	*Eco*R1
LWIN5	5' CC GGG GAT CCG TTT TTA GAT GGA ATA GA 3'	28 bp	*Bam*H1
LWIN3	5' G GGA GCT CGA CTA ATC CTC ATC CTG CT 3'	27 bp	*Sac*1

The plasmid pSVC21 containing the entire HIV-1 genome was used as template and the following PCR reactions were set up:

Reverse transcriptase

5 µl 10X PCR buffer
5 µl 500 µM dNTP-mix
1 µl 5U/µl Taq polymerase
1.5 µl 15 mM LWRT5 primer
1.5 µl 15 mM LWRT3 primer
1 µl 170 pg/µl pSVC21
20 µl ddH_2O
35 µl

Integrase

5 µl 10X PCR buffer
5 µl 500 µM dNTP-mix
1 µl 5U/µl Taq polymerase
1.5 µl 15 mM LWIN5 primer
1.5 µl 15 mM LWIN3 primer
1 µl 170 pg/µl pSVC21
20 µl ddH_2O
35 µl

Two parallel reactions were run for each amplification. The thermal cycling was performed on a Perkin Elmer 2400 termocycler. The PCR cycling program is described in Figure 7.

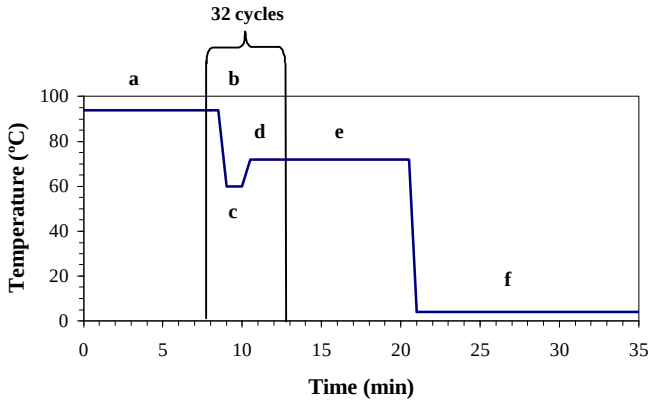


Figure 7. Description of the PCR cycling program. a) 8 minutes of denaturation at 94°C. b) 30 seconds at 94°C. c) 60 seconds annealing starting at 60°C and increasing 0.4°C per cycle. d) 2 minutes extension at 72°C. e) 8 minutes at 72°C. f) Cooling of the samples at 4°C.

2.13.2 Purification of the PCR products

The products from the two parallel PCR reactions were pooled giving approximately 60 µl of PCR product for both IN and RT. 6 µl 3 M NaAc, pH 4.8 was added before vortexing. Following, 132 µl of ice-cold 100 % ethanol was added. The mix was vortexed and left at for 30 minutes at -20°C. The samples were then centrifuged at maximum speed in an eppendorf centrifuge, and the supernatant removed. 1 ml 70 % ethanol was added, and the centrifugation was repeated. After removing the supernatant again, the pellet was dried for 10 minutes in a vacuum centrifuge. The pellet was then dissolved in 30 µl ddH_2O.

2.14 Construction of *in vitro* transcription and translation vectors encoding IN and RT

The two *in vitro* transcription vectors were made by inserting the genes for HIV-1 integrase (IN) and HIV-1 reverse transcriptase (RT) into the *in vitro* transcription and translation vector pCITE-5b(+) from Novagen. The genes were amplified from pSVC21 (containing the entire HIV-1 genome from the HxB2 clone) by PCR (see 2.13).

2.14.1 Digestion with restriction enzymes

The following reactions were performed:

pCITE-5b(+) cut with *Eco*R1 and *Bam*H1

10 µl 0.5 µg/µl pCITE-5b(+)
2 µl 10X Multicore Buffer
1 µl 12 U/µl *Eco*R1
1 µl 10 U/µl *Bam*H1
6 µl ddH_2O
20 µl

pCITE-5b(+) cut with *Sac*1 and *Bam*H1

10 µl 0.5 µg/µl pCITE-5b(+)
2 µl 10X Multicore Buffer
1 µl 10 U/µl *Sac*1
1 µl 10 U/µl *Bam*H1
6 µl ddH_2O
20 µl

The PCR products were incubated with an excess of restriction enzymes:

Reverse transcriptase PCR product cut with *Eco*R1 and *Bam*H1

5 µl reverse transcriptase PCR product
2 µl 10X Multicore Buffer
3 µl 12 U/µl *Eco*R1
3 µl 10 U/µl *Bam*H1
7 µl ddH_2O
20 µl

Integrase PCR product cut with *Sac*1 and *Bam*H1

5 µl integrase PCR product
2 µl 10X Multicore Buffer
3 µl 12 U/µl *Sac*1
3 µl 10 U/µl *Bam*H1
7 µl ddH_2O
20 µl

All four reactions were incubated at 37°C over night and the reactions were stopped by adding 5 µl 6X loading buffer.

2.14.2 Electrophoresis in low meltingpoint (LMP) agarose

A 1 % LMP agarose gel was used. A total amount of 25 µl from the restriction enzyme cutting reaction was applicated on the gel. The gel was run for 1.5 hours at 100 V. The bands were visualised with ethibium bromide in UV- light. The desired bands were cut out from the gel and put in a sterile eppendorf tube.

2.14.3 Ligation in LMP agarose

The bands from the LMP agarose gel-electrophoresis were melted at 70°C for 10 minutes. The following ligation reactions were performed:

Reverse transcriptase

10 µl pCITE-5b(+) cut with *Eco*R1 and *Bam*H1
5 µl RT PCR product cut with *Eco*R1 and *Bam*H1
10 µl ddH_2O
} Kept at 70°C

This was mixed thoroughly by pipetting up and down and

3 µl 10 X Ligation buffer
2 µl T4 DNA ligase
} 37 °C

was added. The reaction was incubated at room temperature over night.

Integrase

10 µl pCITE-5b(+) cut with *Sac*1 and *Bam*H1
5 µl IN PCR product cut with *Sac*1 and *Bam*H1
10 µl ddH_2O
} Kept at 70°C

This was also mixed thoroughly by pipetting up and down and

3 µl 10 X Ligation buffer
2 µl T4 DNA ligase
} 37 °C

was added. This reaction was also incubated at room temperature over night.

After the incubation, the reactions were heated to 70°C and 60 µl 70°C ddH_2O was added to each reaction. 10 µl of this was used to transform 100 µl electro-competent *E.coli* cells. After transformation and identification of positive clones, the plasmids were characterised by restriction enzyme analysis and purified.

2.14.4 Sequencing of constructs

The following primers were used in the sequencing reactions:

Name	Sequence	Length
CITE	5’ GGG GAC GTG GTT TTC CTT TG 3’	20 bp
T7 terminator	5’ GCT AGT TAT TGC TCA GCG G 3’	19 bp

These primers are flanking the multiple cloning sites of the plasmid pCITE-5b(+).
The following reactions were set up:

pCITE-RT forward

4µl Big Dye
2µl 0.5 µg/µl pCITE-RT
1.6 µl 1µM CITE primer
2.4 µl ddH_2O
10 µl

pCITE-RT reverse

4µl Big Dye
2µl 0.5 µg/µl pCITE-RT
1.6 µl 1µM T7 terminator primer
2.4 µl ddH_2O
10 µl

pCITE-IN forward

4µl Big Dye
2µl 0.5 µg/µl pCITE-IN
1.6 µl 1µM CITE primer
2.4 µl ddH_2O
10 µl

pCITE-IN reverse

4µl Big Dye
2µl 0.5 µg/µl pCITE-IN
1.6 µl 1µM T7 terminator primer
2.4 µl ddH_2O
10 µl

The reagents were mixed in PCR tubes and spun down in an eppendorf centrifuge. Then the thermocycling was performed as follows: After an initial denaturing step at 94°C for 5 minutes, 25 cycles with 10 seconds at 96°C, then 5 seconds at 50°C following 4 minutes at 60°C was run.

To precipitate the DNA, 2 µl 3M sodium acetate pH 4.6, 50 µl 95% ethanol and the DNA was mixed, the tubes vortexed and incubated on ice for 10 minutes. After the incubation, the samples were centrifuged at 13000 rpm for 20 minutes. The supernatants were removed with suction and the pellets were rinsed with 250 µl 70% ethanol. The samples were centrifuged again at 13000 rpm, this time for 5 minutes. The ethanol was removed and the pellets dried in a vacuum centrifuge for 15 minutes.

The sequencing was performed on an ABI PRISM 377 DNA Sequencer at the Sequencing Facility at the University of Bergen.

2.15 *In vitro* transcription and translation

2.15.1 Production of reverse transcriptase and integrase *in vitro*

The two proteins were produced *in vitro* with TNT® T7 Quick Coupled Transcription / Translation System from Promega. They were visualised after Western blotting by using Enhanced Chemiluminiscence (ECL).

The following reactions were set up:

Negative control (no DNA added):

40 µl TNT Quick Master Mix
1 µl 1 mM Methionine
9 µl Nuclease free water
50µl

Reverse transcriptase:

40 µl TNT Quick Master Mix
1 µl 1 mM Methionine
2 µl (0,5 µg/µl) pCITE-RT
7 µl Nuclease free water
50 µl

Integrase :

40 µl TNT Quick Master Mix
1 µl 1 mM Methionine
2 µl (0,5 µg/µl) pCITE-IN
7 µl Nuclease free water
50 µl

The reactions were incubated at 30°C for 1 hour, and then kept at –20°C. The proteins were characterised by SDS-PAGE and Western blotting.

2.16 Protein methods

2.16.1 SDS-PAGE

The separating gels were made after the following table:

	8 % gel	**10 % gel**	**12 % gel**	**15 % gel**
ddH_2O	4.55 ml	4.1 ml	3.35 ml	2.4 ml
1.5 M Tris-HCl, pH 8.8	2.5 ml	2.5 ml	2.5 ml	2.5 ml
10 % SDS	0.1 ml	0.1 ml	0.1 ml	0.1 ml
30 % acrylamide, bis	2.8 ml	3.35 ml	4.0 ml	5.0 ml
10 % APS	50 µl	50 µl	50 µl	50 µl
TEMED	5 µl	5 µl	5 µl	5 µl

The APS and the TEMED were added just before use.

Stacking gel:

6.1 ml ddH2O
2.5 ml 0.5 M Tris-HCl, pH 6.8
0.1 ml 10 % SDS
1.3 ml 30 % acrylamide, bis
50 µl APS
10 µl TEMED

The stacking gel solution was added after the separating gel has solidified

The samples were added sample buffer and boiled for 5 minutes at 95°C. They were then cooled on ice before loading on the gel. The gels were run at 100 V until the colour front has reached the bottom of the gel.

2.16.2 Western blotting

The gel and a nitro-cellulose membrane were placed between 4 filter papers and two "pads" and placed in a blotting chamber. The chamber was filled with 1X blotting buffer and placed on ice. The blotting was done at 100 V for 45 minutes. The membrane was then rinsed two times with TBS and incubated in Blocking solution for 1 hour to block the rest of the protein binding sites on the membrane. After the blocking procedure, the membrane was rinsed quickly two times in TBS-Tween (0.05 % Tween 20), then rinsed for 15 minutes and two times 5 minutes. Next the membrane was incubated with a primary antibody over night at 4°C or for 1 hour at room temperature. The membrane was then rinsed quickly two times, then for 15 minutes and two times 5 minutes following 1 hour incubation with a secondary antibody conjugated to POD. The membrane was then washed again quickly three times with TBS-Tween and then for 15 minutes. The last rinsing was done two times 5 minutes with TBS. The membrane was now ready to be developed.

2.16.3 Detection of proteins on Western blots

The detection of proteins on Western blots was done with ECL™ Western blotting detection reagents from Amersham Life Science. After the incubation with secondary antibody conjugated to POD and subsequent rinsing, the membrane was incubated for 1 minute in a mix of 0.5 ml detection solution 1 and 0.5 ml of detection solution 2. Following incubation, the membrane was wrapped in plastic and placed in an X-ray film cassette, protein side up. It

was important to avoid air-bubbles. The rest of the procedure was performed in the dark. A sheet of audioradiography film was placed on top of the plastic-wrapped membrane, and exposed for 10 minutes. The film was replaced with a new one, and the exposed film was developed. The second film was exposed for about 20 hours.

2.17 The ELISA procedure

Two different enzyme-linked immunosorbent assays were performed for detection of RT-IN interaction (modified from Greenway,*et al* [85]).

The first assay involves interaction of RT with immobilised IN, while the second assay involves interaction of IN with immobilised RT. In the first case, 96-well polystyrene microtiter plates were coated with 50 µl of the *in vitro* translation reaction with IN, and with the negative control reaction as a control. Both were diluted in PBS (See Figure 8a for details). The coating was performed over night at 4°C. The wells were then washed three times with PBS-Tween. To block the remaining binding sites, 150 µl of 1% gelatine dissolved in PBS was added to each well, and the plates were incubated for 1 hour at 37°C. After further washing with PBS-Tween, a second coating step was performed. Here 50 µl of the *in vitro* translation reaction with RT diluted in PBS or PBS alone was added to each well (See Figure 8b for details). The plates were incubated for 2 hours at 37°C. Following washing with PBS-Tween, anti-RT antibodies 8C4 and 11B7 [83] diluted 1:100 in PBS were added to each well (50 µl per well), and the plates were incubated for 1.5 hours at 37°C. Wells were again washed as described above and then incubated with anti-mouse Ig conjugated to biotin (diluted 1:1000), 50 µl per well. The incubation was done for 1 hour at 37°C. Following further washing steps, the wells were incubated for 30 minutes at 37°C with streptavidin-HRP diluted 1:1000 in PBS, 50 µl per well. Binding was detected by using *o*-phenylenediamine as the substrate. The reaction was stopped after 15 minutes by adding 100 µl of 1M H_2SO_4. Absorbance was measured by using a plate reader at 492 nm.

Concerning the second case, 96-well polystyrene microtiter plates were coated with 50 µl of the *in vitro* translation reaction with RT, and also here the negative control reaction as a control. Both were diluted in PBS (see Figure 8a for details). The coating was performed over night at 4°C. The wells were then washed three times with PBS-Tween. To block the remaining sites of the wells, 150 µl of 1% gelatine dissolved in PBS was added to each well, and the plates were incubated for 1 hour at 37°C. After further washing with PBS-Tween, a second coating was also here performed. 50 µl of the *in vitro* translation reaction with IN

diluted in PBS or PBS alone was added to each well (see Figure 8b for details). The plates were incubated for 2 hours at 37°C. Following washing with PBS-Tween, anti-IN antibodies 8G4 and 8E5 [84] diluted 1:100 in PBS were added to each well (50 µl per well), and the plates were incubated for 1.5 hours at 37°C. The rest of the procedure was performed as described above.

	A	B	C	D	E	F	G	H	I	J	K	L
1												
2												
3		50 µl 1:2 IN from pCITE-IN	50 µl 1:4 IN from pCITE-IN	50 µl 1:8 IN from pCITE-IN	50 µl 1:16 IN from pCITE-IN	50 µl 1:32 IN from pCITE-IN	50 µl 1:2 RT from pCITE-RT	50 µl 1:4 RT from pCITE-RT	50 µl 1:8 RT from pCITE-RT	50 µl 1:16 RT from pCITE-RT	50 µl 1:32 RT from pCITE-RT	
4		50 µl 1:2 Negative control	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1:2 Negative control	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	
5		50 µl Purified RT 10 µg	50 µl Purified RT 5 µg	50 µl Purified RT 2.5 µg	50 µl Purified RT 1.25 µg	50 µl Purified RT 0.625 µg	50 µl Purified IN 10 µg	50 µl Purified IN 5 µg	50 µl Purified IN 2.5 µg	50 µl Purified IN 1.25 µg	50 µl Purified IN 0.625 µg	
6		50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	
7												
8												

Figure 8a. Overview of the first coating step in the ELISA-procedure. Overview of the proteins added in the first coating.

	A	B	C	D	E	F	G	H	I	J	K	L
1												
2												
3		50 µl 1:2 RT from pCITE-RT	50 µl 1:2 RT from pCITE-RT	50 µl 1:2 RT from pCITE-RT	50 µl 1:2 RT from pCITE-RT	50 µl 1:2 RT from pCITE-RT	50 µl 1:2 IN from pCITE-IN	50 µl 1:2 IN from pCITE-IN	50 µl 1:2 IN from pCITE-IN	50 µl 1:2 IN from pCITE-IN	50 µl 1:2 IN from pCITE-IN	
4		50 µl 1:2 RT from pCITE-RT	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1:2 IN from pCITE-IN	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	
5		50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	
6		50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	50 µl 1 x PBS	
7												
8												

Figure 8b. Overview of the second coating step in the ELISA-procedure. Overview of the proteins added in the second coating.

3 Results

3.1 Production of *in vitro* translation vectors

3.1.1 Amplification of coding regions by PCR

In order to make the HIV-1 proteins RT and IN by *in vitro* transcription and translation, the *in vitro* translation vector pCITE-5b(+) was chosen. The first step in this process was to amplify the coding regions of the RT and IN genes. These regions were amplified by PCR from the plasmid pSVC21 containing the complete HIV-1 HxB2 genome. Since there were no appropriate cleavage sites for restriction enzymes flanking the coding regions, new ones had to be designed in the primers. A cleavage site for *Bam*HI was introduced in the 5' end for both the genes. For RT, an *Eco*RI site was introduced in the 3'end. These sites enable in-frame cloning of RT into the vector pCITE-5b(+). IN has a cleavage site for *Eco*RI in the coding region of the gene, therefore a cleavage site for *Sac*I was chosen. With these cleavage sites, also IN can be cloned in the correct reading frame into the vector pCITE-5b(+).

Two identical PCR reactions were run for each primer pair, as described in Methods (2.13). After the PCR-reactions, 2 µl of PCR product was run on a 1% agarose gel
(Figure 9).

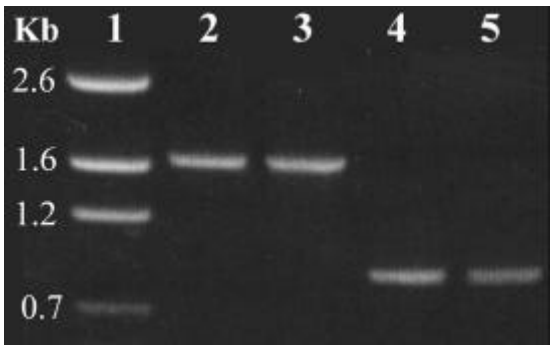


Figure 9. Characterisation of PCR products on a 1% agarose gel. The coding regions of RT and IN was amplified by PCR from the plasmid pSVC21 containing the complete HIV-1 HxB2 genome. 1) pGEM standard. 2) and 3) PCR product with RT primers from two identical PCR reactions. 4) and 5) PCR product with IN primers from two identical PCR reactions.

To estimate the size of the PCR products, the path length of the pGEM standard was measured and plotted against the number of basepairs (Table 1)

Table 1. Path length and corresponding size of the pGEM standard

Path length	Size (bp)
1 cm	2645
2.1 cm	1605
2.75 cm	1199
4 cm	676

From this a standard curve was drawn and the sizes of the two PCR products estimated (Figure 10).

The coding region of RT is 1680 basepairs, a size that correspond well with the estimation of 1800 basepairs from the standard curve. For IN the coding region is 891 basepairs, which also corresponds to the estimation of 825 basepairs.

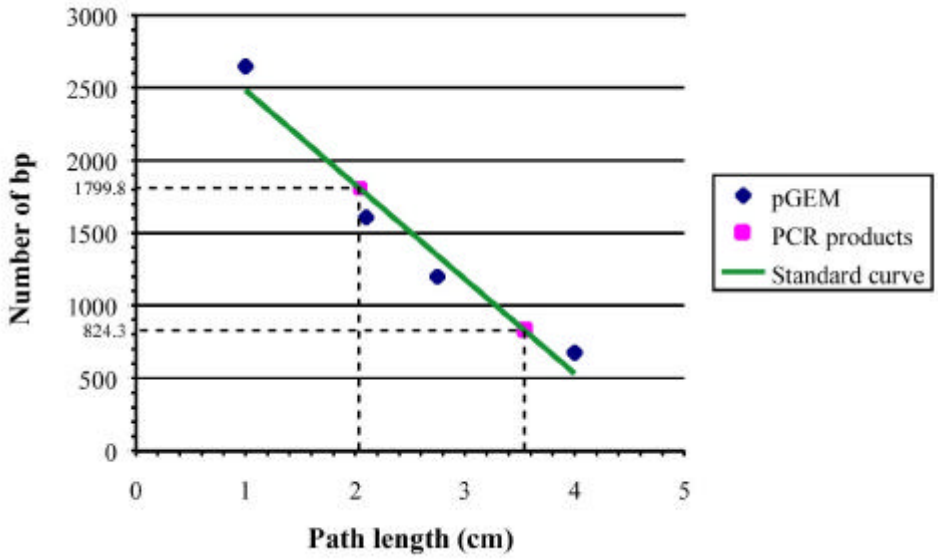


Figure 10. Standard curve for the estimation of fragment size of PCR products. The path length of the pGEM standard on an 1% agarose gel was measured and plotted against the size of the bands. From this a standard curve was drawn. The path length of the two PCR products was measured, and the sizes estimated from the standard curve.

The RT and IN coding region PCR products were purified to get a high concentration, and cut with the appropriate restriction enzymes before cloning them into the vector pCITE-5b(+).

3.1.2 Cloning of RT and IN into the vector pCITE-5b(+)

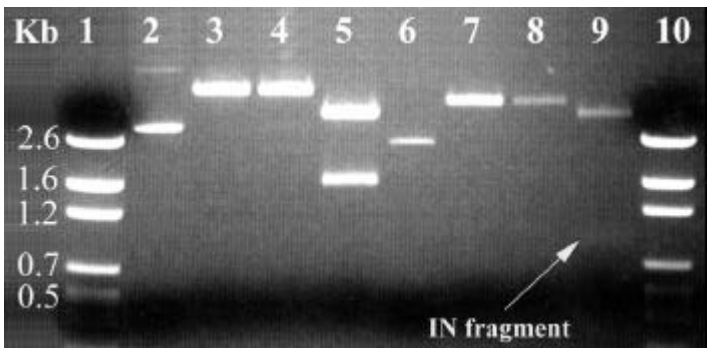
To be able to clone the amplified coding regions of RT and IN into the vector pCITE-5b(+), the vector first had to be cut with the appropriate restriction enzymes. Then the cloning was performed as described in Methods (2.14). Electrocompetent bacteria were transformed with the cloning products and plasmids were isolated from the positive colonies. 10 µl of an 1:100 dilution of the resulting product was cut with restriction enzymes and run on an 1% agarose gel as described in Table 2 and Figure 11. The anticipated restriction enzyme digest pattern, made with the computer program Gene Construction Kit, is also shown.

The plasmid containing the RT gene is called pCITE-RT, and the plasmid containing the IN gene is called pCITE-IN.

Table 2. Restriction enzyme analysis of pCITE-RT and pCITE-IN

Lane	Description
1	pGEM standard
2	Uncut pCITE-RT
3	pCITE-RT cut with *Bam*HI
4	pCITE-RT cut with *Eco*RI
5	pCITE-RT cut with *Bam*HI and *Eco*RI
6	Uncut pCITE-IN
7	pCITE-IN cut with *Bam*HI
8	pCITE-IN cut with *Sac*I
9	pCITE-RT cut with *Bam*HI and *Sac*I
10	pGEM standard

a)



b)

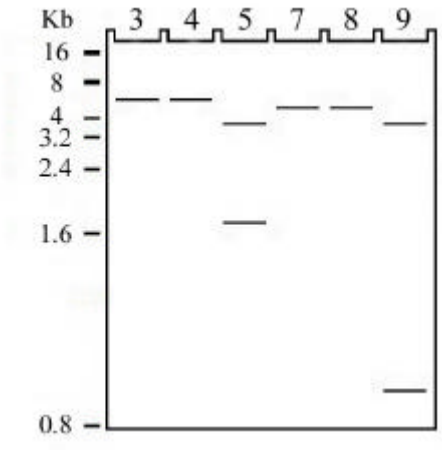


Figure 11. Restriction enzyme analysis of pCITE-RT and pCITE-IN. a) The plasmids are cut with restriction enzymes and run on an 1% agarose gel. 1) pGEM standard. 2) Uncut pCITE-RT. 3) pCITE-RT cut with *Bam*HI. 4) pCITE-RT cut with *Eco*RI. 5) pCITE-RT cut with both *Bam*HI and *Eco*RI. 6) Uncut pCITE-IN. 7) pCITE-IN cut with *Bam*HI. 8) pCITE-IN cut with *Sac*I. 9) pCITE-RT cut with both *Bam*HI and *Sac*I. 10) pGEM standard. b) Anticipated restriction enzyme digest pattern created with the computer program Gene Construction Kit. The numbering is the same as in a).

In lane 2 the strong band corresponds to supercoiled plasmid pCITE-RT. The weaker band above is open circular plasmid. Lanes 3 and 4 have only linear plasmid as expected. In lane 5 a fragment of approximately 1700 bp is clearly visible. This corresponds to the RT coding region cut out from the plasmid pCITE-RT. The other band is the linear vector. In lanes 6-9, the plasmid pCITE-IN is used. This plasmid has a slightly lower concentration, resulting in weaker bands. In lane 6 the band is supercoiled pCITE-IN. In lanes 7 and 8 the bands corresponds to linear plasmid. In lane 9 the vector without the IN fragment is visible. The band corresponding to the IN fragment is hardly visible on the picture (see arrow), but was visible when the gel was examined under UV-light. This corresponds well with the pattern obtained when the plasmids were digested using the computer program Gene Construction Kit (Figure 11b).

These results strongly suggests that the genes HIV-1 RT and IN was successfully inserted into the vector pCITE-5b(+). The plasmids thus obtained are illustrated in Figure 12.

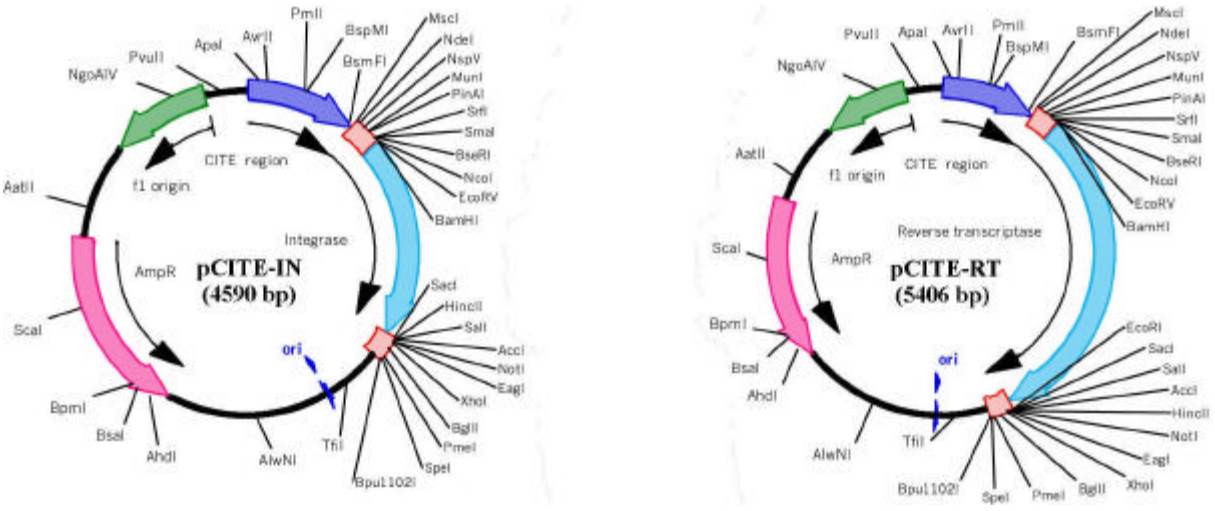


Figure 12. The plasmids pCITE-IN and pCITE-RT. The plasmid maps are made using Gene Construction Kit.

3.1.3 Determination of the concentrations of pCITE-RT and pCITE-IN

The plasmids were purified using a Qiagen Megaprep protocol (see 2.12.3). The purified plasmids were then subjected to a spectrophotometric analysis to determine the concentrations (Table 3).

Table 3. Determination of plasmid concentrations

Plasmid	Dilution	A_{260}	Concentration
pCITE-RT	1:100	0.491	2.5 µg/µl
pCITE-RT	1:1000	0.046	2.3 µg/µl
pCITE-IN	1:100	0.273	1.4 µg/µl
pCITE-IN	1:1000	0.035	1.75 µg/µl

For pCITE-RT the difference in concentration from the two measurements is small, the conclusion is therefore that the concentration is about 2.4 µg/µl. The difference between the measurements for pCITE-IN is bigger, but since the A_{260} value for the 1:1000 dilution is so low, a concentration of 1.4 µg/µl is believed to be the most correct.

3.1.4 Sequence analysis of recombinant constructs

The plasmids were sequenced using both forward and reverse primers as described in Methods (2.14.4). 3-400 nucleotides could easily be read for both RT and IN. The chromatograms obtained was visualised using the computer program Chromas 1.45 (Figure 13). The computer program ClustalX was used to compare the sequences with the published HIV-1 HxB2 sequence. The only differences found was where an N had been placed in the sequence. By carefully examine the chromatogram the right nucleotides could be determined. Only the startpoints and the endpoints of the genes are visualised. The RT gene ends with an N, but a red peak is clearly visible on the chromatogram. This strongly suggests that the right nucleotide is a T, which corresponds with the HIV-1 HxB2 sequence.

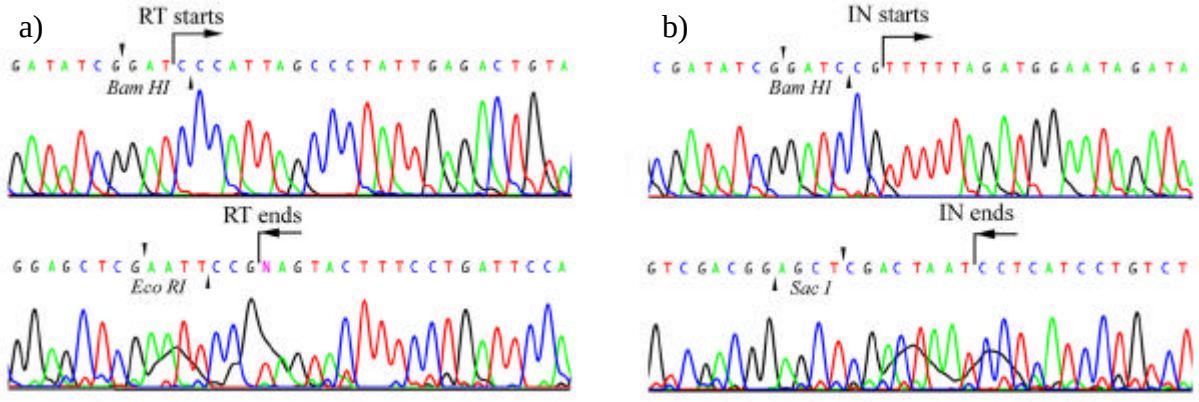


Figure 13. Chromatograms from the four sequencing reactions. a) The start- and endpoint of the RT-gene in the plasmid pCITE-RT is visualised. The restriction enzyme sites are also shown. b) The start- and endpoints of IN in the plasmid pCITE-IN with restriction enzyme sites. Chromas 1.45 is used to visualise the chromatograms.

3.2 Description of the anticipated proteins

Since both RT and IN are expressed as a polyprotein (Pol) in HIV-1, neither of the proteins have the start codon needed for translation. These are therefor provided in the vector, yielding slightly longer polypeptide chains then for the native proteins. RT is in the N-terminal part of Pol, and thus has no stop-codon either. This is also provided in the vector, giving RT from pCITE-RT with both an N-terminal and a C-terminal extension. IN has a stop-codon, since IN is the C-terminal part of Pol.

By using the computer sequence analysis program ClustalX, the anticipated RT and IN sequences was compared to the HxB2 Pol sequence. The results are shown in Figure 14.

a)

```
CLUSTAL X (1.64b) multiple sequence alignment

HXB2            ----------------------------------------------PISPIETVPVKLKPG 15
PCITE-RT        MATTHMKETAAAKFERQHMDSPIDGTGSDDDDKSPGFSSTMAISDPISPIETVPVKLKPG 60
                                                              ***************

HXB2            MDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLV 75
PCITE-RT        MDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLV 120
                ************************************************************

HXB2            DFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSI 135
PCITE-RT        DFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSI 180
                ************************************************************

HXB2            NNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEI 195
PCITE-RT        NNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEI 240
                ************************************************************

HXB2            GQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVN 255
PCITE-RT        GQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVN 300
                ************************************************************

HXB2            DIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAELELAENREILKEPVH 315
PCITE-RT        DIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAELELAENREILKEPVH 360
                ************************************************************

HXB2            GVYYDPSKDLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKI 375
PCITE-RT        GVYYDPSKDLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKI 420
                ************************************************************

HXB2            TTESIVIWGKTPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIV 435
PCITE-RT        TTESIVIWGKTPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIV 480
                ************************************************************

HXB2            GAETFYVDGAANRETKLGKAGYVTNRGRQKVVTLTDTTNQKTELQAIYLALQDSGLEVNI 495
PCITE-RT        GAETFYVDGAANRETKLGKAGYVTNRGRQKVVTLTDTTNQKTELQAIYLALQDSGLEVNI 540
                ************************************************************

HXB2            VTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAG 555
PCITE-RT        VTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAG 600
                ************************************************************

HXB2            IRKVL-------------------- 560
PCITE-RT        IRKVLRNSSSVDKLAAALEHHHHHH 625
                *****
```

b)

```
CLUSTAL X (1.64b) multiple sequence alignment


HxB2            -------------------------------------------FLDGIDKAQDEHEK 14
pCITE-IN        MATTHMKETAAAKFERQHMDSPIDGTGSDDDDKSPGFSSTMAISDPFLDGIDKAQDEHEK 60
                                                           **************

HxB2            YHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVIL 74
pCITE-IN        YHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVIL 120
                ************************************************************

HxB2            VAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTGATVRAACWWAG 134
pCITE-IN        VAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTGATVRAACWWAG 180
                ************************************************************

HxB2            IKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGY 194
pCITE-IN        IKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGY 240
                ************************************************************

HxB2            SAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRNPLWKGPAKLLWKGEGAVVIQDN 254
pCITE-IN        SAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRNPLWKGPAKLLWKGEGAVVIQDN 300
                ************************************************************

HxB2            SDIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED 288
pCITE-IN        SDIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED 334
                **********************************
```

Figure 14. Sequence alignment. a) Alignment of the HIV-1 HxB2 RT sequence with the anticipated protein from the pCITE-RT plasmid. b) Alignment of the HIV-1 HxB2 IN sequence with the anticipated protein from the pCITE-IN plasmid.

This shows that the RT protein translated from the plasmid pCITE-RT will have a 45 aa N-terminal extension and a 20 aa C-terminal extension. The N-terminal extension contains an S-tag; which can be utilised to purify the protein. In the C-terminal part, a His-tag has been added. This tag can also be used as a means to purify the protein. The IN protein produced from pCITE-IN will only have the N-terminal extension, since it already contains a stop-codon. This N-terminal extension is the same as for the RT protein.

Since both the proteins produced will have additional amino acids, their molecular weights will be different from the native viral proteins. To calculate the expected molecular weights, the ProtParam tool from the ExPASy (Expert Protein Analysis System) proteomics server of the Swiss Institute of Bioinformatics (SIB) was used. The results obtained when using the translated sequences from pCITE-RT and pCITE-IN are given in Table 4.

Table 4. Anticipated molecular weights of RT and IN from pCITE-RT and pCITE-IN.

Native viral proteins		**pCITE proteins**	
IN	31 kDa	IN	37 kDa
RT (p66)	66 kDa	RT (p66)	71.5 kDa
RT (p51)	51 kDa	RT (p51)	56 kDa

These results are important for the interpretation of results from i. e. Western blots, where the bands will appear higher up on the blot than expected from molecular weights of the native proteins.

3.3 *In vitro* transcription and translation

3.3.1 Production of protein batches

The proteins produced by *in vitro* transcription and translation was to be used in a protein-protein interaction study with ELISA. For this purpose, 50 µl of protein from one *in vitro* transcription and translation reaction was not enough. Therefore nine parallel reactions were set up for both RT and IN. The procedure was followed as described in Methods (2.15.1). After the incubation, a total amount of 50 µl of reticulocyte lysat containing either RT or IN was obtained from each reaction. The concentration of RT and IN in solution was not determined at this point.

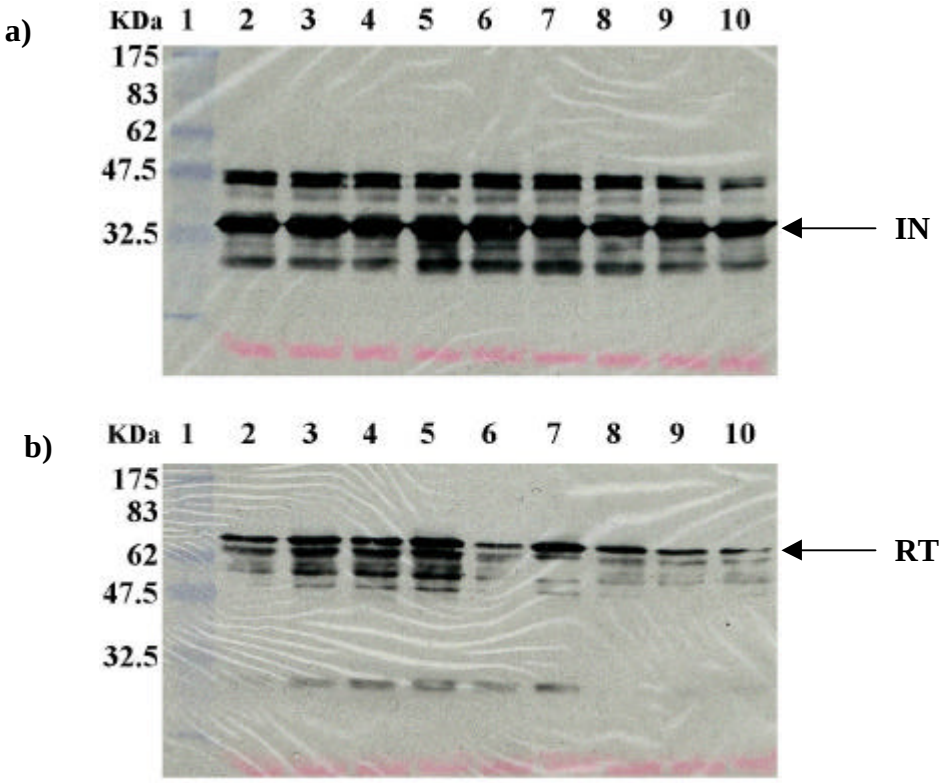


Figure 15. *In vitro* translated HIV-1 RT and IN. a) 5 µl of 9 parallel *in vitro* transcription and translation reactions with pCITE-IN was run on a 10 % SDS-acrylamide gel and blotted onto a nitrocellulose membrane. IN was visualised using the IN-antibodies 8G4 and 8E5 [84], anti-mouse-POD as secondary antibody and detected by ECL. b) 5 µl of 9 parallel *in vitro* transcription and translation reactions with pCITE-RT was run on a 10 % acrylamide gel and blotted onto a nitrocellulose membrane. RT was visualised using the RT-antibodies 8C4 and 11B7 [83], anti-mouse-POD as secondary antibody and detected by ECL.

Following, 5 µl from each reaction was run on an 10 % acrylamide gel and visualised by Western blotting using a mix of three different monoclonal antibodies for each protein (Figure 15).

In Figure 15a, the nine parallel *in vitro* translation reactions producing IN are shown. A strong band above 32.5 kDa is clearly visible for all nine reactions. This is in good accordance to the results obtained using the ProtParam tool from the ExPASy proteomics server of SIB (see Table 4). The results therefor strongly suggest that IN has been produced. There are in addition several other bands visible. There are two strong bands at approximately 47.5 kDa and a weaker band below 30 kDa. Possible explanations for these bands are given in Discussion (4.2.2). In Figure 15b, the nine *in vitro* translation reactions for RT are shown. On this blot, a band between 62 and 83 kDa is visible. Also this is in good accordance to the results obtained using the ProtParam tool from the ExPASy proteomics server of SIB (see Table 4) suggesting that RT as well has been succesfully produced. A strong lower band is visible on this blot, in addition to several other weaker bands. See Discussion (4.2.2) for further details.

The reason for the difference in intensity of the RT and IN bands need not be a difference in concentration. The intensity of bands with the same concentration varies from time to time when ECL is used, due to the high sensitivity of this system.

After the verification of the protein products, the nine *in vitro* translation reactions for each protein were pooled, giving a batch of protein to use in the ELISA-experiments. Before the ELISA-experiments can be carried out, it is thus important that the protein concentrations are determined.

3.3.2 Determination of protein concentrations

Before assaying for RT-IN interactions, an approximately protein concentration must be decided. Since the proteins produced are in a reticulocyte lysate containing many other proteins, regular protein concentration assays can not be utilised. The protein concentrations are therefore approximated on Western blots. 5 µl of the lysates containing either IN or RT was run on an acrylamide gel together with from 0.1 to 1 µg of purified IN or RT. The results are shown in Figure 16.

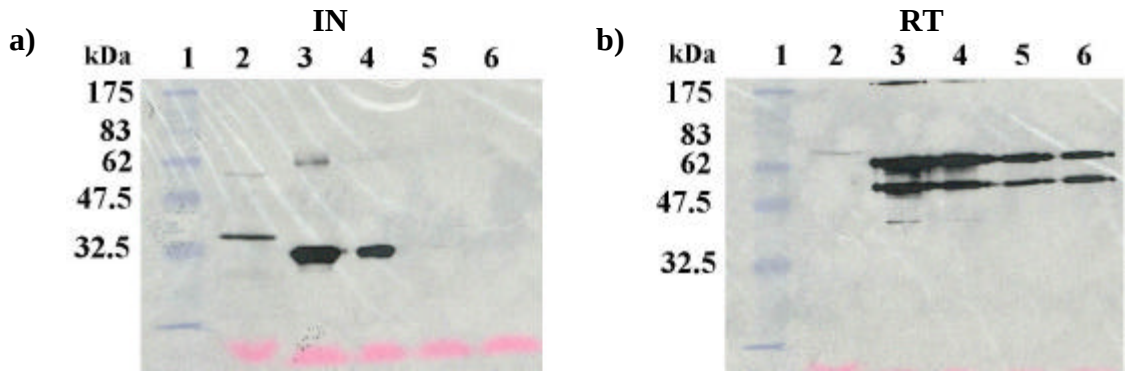


Figure 16. Determination of protein concentrations. a) The samples were run on a 10% SDS-acrylamide gel and blotted onto a nitrocellulose membrane. 1) 15 µl Broad-Range protein standard. 2) 5 µl of the IN pool. 3) 1 µg of purified IN. 4) 0.5 µg of purified IN. 5) 0.2 µg of purified IN. 6) 0.1 µg of purified IN. b) The samples were run on a 10% acrylamide gel and blotted onto a nitrocellulose membrane.1) 15 µl Broad-Range protein standard. 2) 5 µl of the RT pool. 3) 1 µg of purified RT. 4) 0.5 µg of purified RT. 5) 0.2 µg of purified RT. 6) 0.1 µg of purified RT.

The band for the *in vitro* translated IN is quite strong. For the purified IN, only the bands corresponding to 5 and 2.5 µg of IN are visible on the blot. The band in lane 2 is weaker than these bands, giving a concentration of between 0.5 and 0.2 µg of IN in the 5 µl sample. This is only an approximation, since it is difficult to determine protein concentrations on Western blots. For RT the situation is different. Here, the band for the *in vitro* translated RT is quite weak, while all the bands for the purified RT are all strong. This means that the amount of *in vitro* translated RT on the blot are less then 0.1 µg from the 5 µl sample.

The concentrations found differ slightly for the two proteins. This can be associated with the difference in size for RT and IN. Nevertheless, the amounts of protein produced should be sufficient for protein-protein interaction studies to be carried out.

3.4 Results with ELISA

In order to detect any physical interactions between RT and IN, a direct ELISA was performed (see 2.17). A standard curve for each protein was first determined, using purified proteins. The results are shown in Table 5 and Figure 17.

Table 5. Standard curves for RT and IN

RT concentration	Absorbance at 492 nm	IN concentration	Absorbance at 492 nm
10 µg	2.239	10 µg	2.059
5 µg	2.056	5 µg	1.991
2.5 µg	2.071	2.5 µg	1.892
1.25 µg	1.905	1.25 µg	1.512
0.625 µg	1.431	0.625 µg	0.824

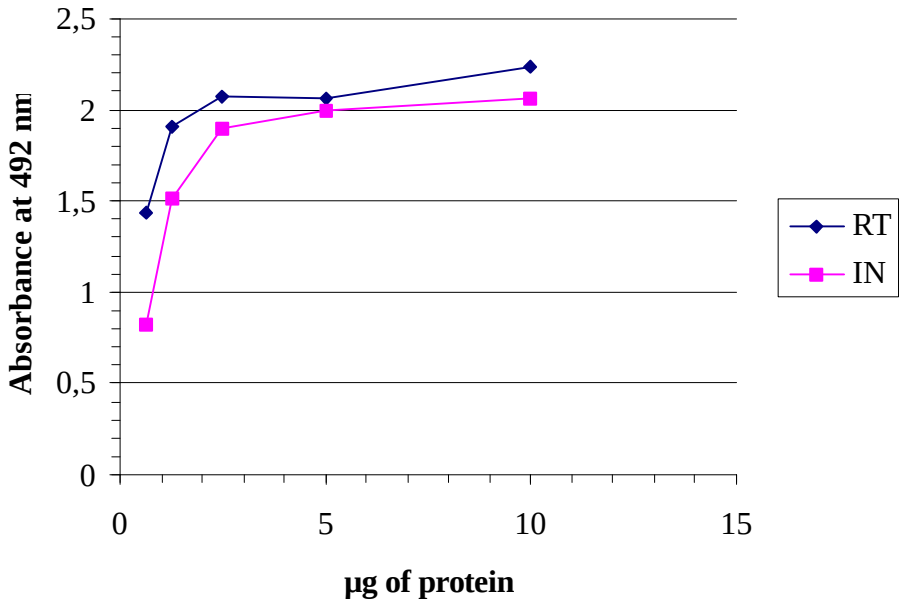


Figure 17. Standard curves for RT and IN. Increasing amounts of purified RT or IN was added to each well, and the procedure followed as described in 2.17

The curves show a saturation of proteins in some of the wells. It is obvious that 10 µg of protein is too much, it seems that also 5 µg is on the upper edge. When the two highest protein concentrations are removed, the standard curves become almost linear (see Figure 18).

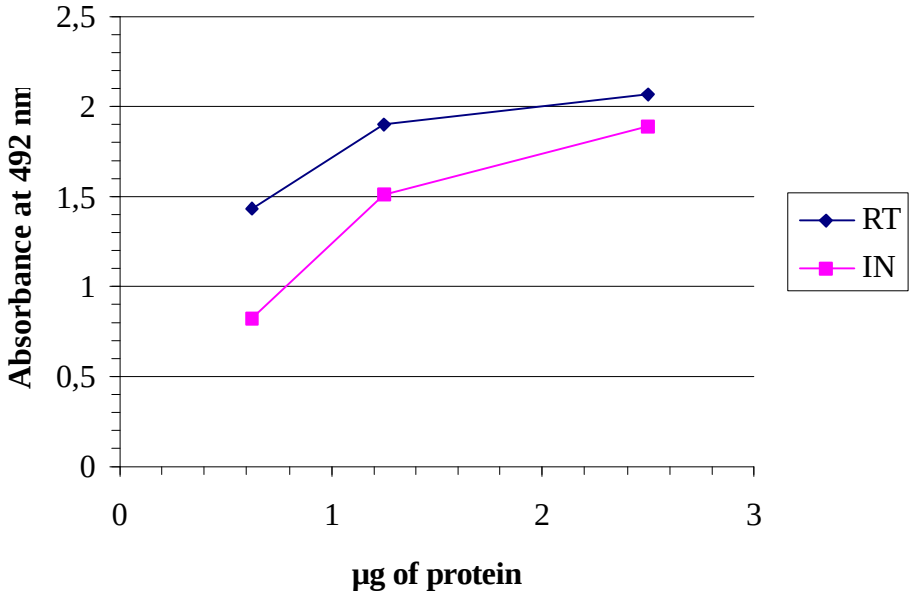


Figure 18. Revised standard curves for RT and IN. The wells where a too high amount of protein was added have been omitted in order to get linear standard curves.

Two different formats of the direct ELISA were performed. In the first format, decreasing amounts of *in vitro* translated IN was added to the wells. Following incubation and blocking, a fixed amount of in *vitro* translated RT was added to all the wells containing IN. A well incubated with the negative control lysate was also added RT. The rest of the procedure was followed as described in 2.17. The second format was performed exactly as the first, only that the first incubation was with RT and the second with IN. The results are shown in Table 6.

Table 6. Results from the ELISA experiments.

Amount of putative interacting proteins added	Abs. 492 nm	Amount of putative interacting proteins added	Abs. 492 nm
50 µl 1:2 IN from pCITE-IN + 50 µl 1:2 RT from pCITE-RT	0.628	50 µl 1:2 RT from pCITE-RT + 50 µl 1:2 IN from pCITE-IN	1.994
50 µl 1:4 IN from pCITE-IN + 50 µl 1:2 RT from pCITE-RT	0.642	50 µl 1:4 RT from pCITE-RT + 50 µl 1:2 IN from pCITE-IN	1.939
50 µl 1:8 IN from pCITE-IN + 50 µl 1:2 RT from pCITE-RT	0.664	50 µl 1:8 RT from pCITE-RT + 50 µl 1:2 IN from pCITE-IN	1.962
50 µl 1:16 IN from pCITE-IN + 50 µl 1:2 RT from pCITE-RT	0.626	50 µl 1:16 RT from pCITE-RT + 50 µl 1:2 IN from pCITE-IN	1.975
50 µl 1:32 IN from pCITE-IN + 50 µl 1:2 RT from pCITE-RT	0.583	50 µl 1:32 RT from pCITE-RT + 50 µl 1:2 IN from pCITE-IN	2.056
50 µl 1:2 Negative control + 50 µl 1:2 RT from pCITE-RT	0.665	50 µl 1:2 Negative control + 50 µl 1:2 IN from pCITE-IN	2.005

As seen in Table 6, the results for both the experiments are the same for all the dilutions and also for the negative controls. This means that the proteins added in the second coating have bound equally to all the wells, regardless of the amount or identity of the proteins added in the first coating. An other conclusion that can be drawn, is that the concentration of the *in vitro* translated IN is much higher than the *in vitro* translated RT. This corresponds well with the results obtained in 3.3.2. From the standard curve (Figure 18), the amount of RT in 50 µl of lysate is less than 0.625 µg. The corresponding amount of IN is between 2.5 and 10 µg. It is impossible to determine this exactly, because of the saturation of protein in the standard curve for IN.

4 Discussion

The HIV-1 proteins reverse transcriptase and integrase have been produced by *in vitro* transcription and translation. The proteins produced have higher molecular weights than the native viral proteins. An ELISA-format has been employed in order to detect possible direct interactions between the synthesised proteins. A protein-protein interaction could not be detected by this method. A number of different parameters could have caused the detection of RT-IN interactions to fail. These parameters need to be changed in future experiments.

4.1 Construction and verification of plasmids containing RT and IN

DNA encoding reverse transcriptase and integrase from the HIV-1 were amplified from a plasmid containing the full-length HxB2 sequence by PCR. This produced fragments with sizes corresponding to the coding regions of both RT and IN as verified by agarose gel electrophoresis. The DNA was purified, cut with restriction enzymes and cloned into the *in vitro* translation vector pCITE-5b(+).

4.1.1 The vector pCITE-5b(+)

This vector was chosen because it is designed for increased translational performance of RNA prepared from cloned DNA sequences. The plasmid carry a segment of the encephalomyocarditis virus (EMCV) RNA 5' non-coding region, that functions as an internal entry point for initiation of translation by mammalian ribosomes. This Cap-Independent Translation Enhancer (CITE) should dramatically increase the *in vitro* translation efficiency of synthetic RNA by rabbit reticulocyte lysates. Since this was the only vector used, it is difficult to say if this vector really gave the best efficiency of translation.

4.1.2 Cloning and sequencing

The cloning was performed directly in low melting-point agarose in order to avoid the extra purification step. The cloning efficiency was high, indicating that the cloning-procedure is efficient as well as being easy to perform. After the cloning, the plasmids were subjected to restriction enzyme analysis in order to verify that RT and IN was inserted into the vector. The results of this analysis confirmed that the right genes were inserted. The sequences of the

plasmids obtained were then determined. This confirmed that the genes were inserted in the right reading frame.

4.1.3 Anticipated proteins

Both RT and IN are enzymatically cleaved from the viral polyprotein Pol when HIV-1 infects host cells. For RT a homodimer of p66 is formed before PR cleaves one of the monomers giving the p66/p51 heterodimer. IN is cleaved from the C-terminal part of Pol giving a protein of 31 kDa. When RT is going to be *in vitro* translated, both start- and stop-codons need to be introduced from the vector sequence. The *in vitro* translation vector pCITE-5b(+) has start-codons and stop-codons flanking the multiple cloning-site in all three reading-frames. Since the RT-gene was cloned into pCITE-5b(+) in the correct reading frame, the start- and stop-codons were introduced. After the start-codon in pCITE-5b(+), an S-tag is present. This will facilitate an eventual purification of the protein using an affinity column. In the C-terminal part, a His-tag is added before the stop-codon. This can also be used as a means to purify the protein. IN is the C-terminal part of Pol, and thus has a stop-codon. The C-terminal His-tag is therefor not introduced here. Only the N-terminal S-tag is added, as for the RT protein.

When these extra amino acids are added to the proteins, their molecular weights change (3.2). This can affect their ability to interact. Additional amino acids can dramatically change the folding of a protein, giving a different 3-dimentional structure. Protein-protein interactions are dependent on tertiary structure, and the effects of the additional amino acids in the *in vitro* translated proteins are not known. An easy way to check if the structure is affected, is by using activity assays for both the proteins. If the proteins are active, their folding is most likely correct and a possible interaction would not be affected. This however, requires purified proteins and was not done in this project.

4.2 *In vitro* transcription and translation

The *in vitro* transcription and translation was performed using the TNT® Quick Coupled Transcription/Translation Systems from Promega. The incubation time was varied from 0.5 to 1.5 hours without detecting any difference neither in the amount nor quality of the protein produced. This is the time the manufacturer recommends for optimal *in vitro* transcription and translation.

4.2.1 Labelling of proteins

At first the proteins were labelled with biotin by the way of biotin-lysyl-tRNAs (Figure 19). This enabled an easy detection on Western blots using streptavidin. However, the effect of biotin on the putative protein-protein interactions between RT and IN is uncertain. Since 9.9% of the residues in RT produced from pCITE-RT are lysines and the corresponding number for IN is 8.4%, a substantial part of the protein would be marked. Lysine residues are usually exposed at the aqueous-facing exterior of the protein. The presence of biotinylated lysines may or may not affect the function of the modified protein. For example, the enzymatic activity of ß-galactosidase is reduced by the incorporation of biotinylated lysines but luciferase and chloramphenicol acetyltransferase (CAT) are affected less [86]. Since the effects on RT and IN are unknown, it was decided to synthesise the proteins unmarked.

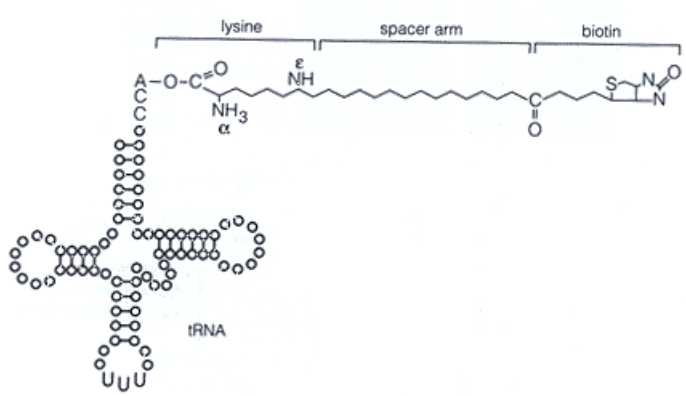


Figure 19. Structure of Transcend™ tRNA. The Transcend™ tRNA is a biotinylated lysyl-tRNA, where the biotin moiety is linked to lysine by a spacer arm. The figure is adapted from Promegas Technical Bulletin No. 182, page 2.

The unmarked proteins were analysed by SDS-PAGE. Several attempts were done to visualise the proteins by Coomasssie-staining. This was never successful; the only bands appearing were the lysate proteins as compared to the negative controls. Since nothing appeared to be on the gel, the procedure for *in vitro* transcription and translation was questioned. Several adjustments of time and temerature of incubation were made, but nothing did appear on the gel when using the Coomassie-staining. Since adjusting the procedure did not yield any changes, an explanation for the missing proteins could be that the concentrations were two low to be detected by Coomassie-staining. To verify this, Western blotting was performed. When using a mix of three different monoclonal antibodies for each protein, bands at the appropriate sizes were present after developing the radiography film. Western blotting is a more sensitive method than Coomassie-staining, and lower protein concentrations can be detected. From 30 to 100 ng of protein per band is detectable by Coomassie staining methods

[87]. With the ECL Western blotting detection reagents, less then 1 pg of protein can be detected.

4.2.2 Production of proteins

At first, nine parallel *in vitro* translation reactions were performed for RT and IN. This was done in order to get a batch of protein that could be used in later ELISA experiments. The resulting proteins were visualised by Western blotting. Bands at the right molecular weights appeared on the blots after developing with ECL (Figure 15). There were also several additional bands. For IN two bands at approximately 47.5 kDa were visible. These bands are too low to be a dimer of IN. A possible explanation is that the bands represent different forms of degraded dimer. There are also several lower bands, which also can represent degradation products of IN. Two different monoclonal antibodies recognising different epitopes of IN were utilised for detection on Western blots. These antibodies show no detectable cross-reactivity with bacterial proteins [84]. It is not known whether there exists cross-reactivity between the antibodies used and proteins in the reticulocyte lysate. For RT, the additional bands were lower than the band corresponding to the produced RT. These bands can correspond to p66 and p51 without one or both of the tags added from the vector. If the RT produced folds correctly, the tags will most likely protrude from the structure and be accessible for proteolytic attacks. An other explanation for the extra bands is that RT is degraded. The antibodies utilised for detection recognises the same epitope on both p66 and p51. These antibodies are not known to cross-react.

The sensitivity of the ECL system is also very high. Small differences in the time of exposure to the radiography film gave big differences in the appearance of the blot. The first bands appearing on the film were still always the bands corresponding to RT and IN.

These nine in *vitro* translation reactions were pooled and a sample from the pool was blotted together with purified RT and IN respectively. This was done in order to get an impression of the concentration of proteins of interest in the pool. For RT a concentration of less than 20 ng/µl was found. The corresponding concentration of IN is between 40 and 100 ng/µl. These concentrations are higher than what could be expected. Promega states that although expression levels for different genetic constructs can vary greatly due to many factors, most gene constructs will generate 50-300 ng of protein in a standard 50 µl reaction in one hour. The high yield of protein using the vectors pCITE-RT and pCITE-IN is probably due to the CITE-sequence in the vector.

4.3 ELISA

4.3.1 Direct ELISA

A direct ELISA-format was employed in order to detect possible protein-protein interactions between *in vitro* translated HIV-1 RT and IN. This format is previously successfully utilised to detect interactions between HIV-1 Nef and the two proteins Lck and MAPK (Mitogen-activated protein kinase) by using purified proteins [85]. In this thesis reticulocyte lysates containing the putative interacting proteins were employed in the ELISA. This resulted in the same results for both the dilution series and the negative control (3.4). One reason for this can be that the blocking was incomplete, but the absorbance for the wells that were only added PBS is so low that this explanation is unlikely. An other explanation can be that there are some endogenous proteins in the reticulocyte lysate that are interacting with RT and IN. This would explain the results for the negative controls. Considering the dilution series, a possible explanation for the identical results for all dilutions could be that the amount of endogenous proteins in the reticulocyte lysate is so high that even a 1:32 dilution is enough to get a saturation of proteins in the wells. Is seems therefor that this ELISA format has to be used with purified proteins in order to avoid unspecific binding.

From the results obtained with ELISA, the concentrations of produced proteins can be verified. From the standard curves a concentration of less that 25 ng/µl is determined for RT. The concentration of IN is more difficult to determine, because the absorbance is at the saturation level. A concentration of less than 100 ng/µl is the most accurate result possible. These concentrations correspond well to those found by Western blotting.

4.3.2 Other ELISA formats

If the reticulocyte lysate is going to be used with ELISA, an indirect ELISA might be better. In an indirect ELISA, the wells are coated with an antibody, not an antigen. Then only the protein of interest will bind to the wells. For the coating with the putative interacting protein however, the lysate has to be used. It is therefore not certain that performing an indirect ELISA will rule out the problem of unspecific binding.

If ELISA is going to be used to detect protein-protein interactions between two *in vitro* translated protein; the proteins should first be purified. This should be done to avoid any non-specific binding between the proteins of interest and the proteins in the reticulocyte lysate.

The concentration of protein would also be higher if the proteins are purified from the reticulocyte lysate.

4.4 RT-IN interactions

It is not evident that direct physical interactions between RT and IN should be present. The PIC consists of several other proteins, both viral and cellular. In addition; both viral RNA and DNA are associated. An interaction between RT and IN may require several other factors linking the two proteins together. Two different methods have previously been used in the search for RT-IN interactions; the yeast two-hybrid system and an *in vitro* binding assay. When the yeast two-hybrid system was applied, no interactions or even hybrid proteins could be detected [82]. The plasmids were successfully rescued from the yeast-cells after transformation, but the two hybrid proteins were undetectable. The most likely explanation for this was that the expression level of proteins was too low. Several factors need to be adjusted if the yeast two-hybrid system is going to be utilised to detect RT-IN interactions.

When an *in vitro* binding assay was applied [75], a GST-IN fusion protein was shown to efficiently pull down a RT heterodimer protein from crude bacterial lysates. Empty G-beads or GST protein-bound G-beads were not able to pull down RT, indicating the specificity of the RT-IN interaction. The RT-IN reaction mixture was pretreated with micrococcal nuclease to rule out the possibility that nucleic acids facilitated the association of RT and IN. This treatment did not decrease the amount of RT pulled down.

It is impossible to conclude if the RT-IN interaction really exists on the basis of only these two publications. With the yeast two-hybrid system no interactions could be detected due to low expression of proteins. With the *in vitro* binding assay, interactions were detected between the two proteins. This was however only done *in vitro*, so how this will be *in vivo* is not yet determined.

4.5 Conclusions

The two *in vitro* translation vectors containing HIV-1 RT and IN seemed to function satisfying. Since the plasmids contained the CITE-sequence, the efficiency of translation was high. The two proteins were not purified, in order to check if interactions between them could be detected without this step. This was not successful, at least not with the ELISA format

utilised in this project. A high degree of unspecific binding was found and a conclusion of interactions could not be drawn.

4.6 Future strategies

The main purpose of identifying RT-IN interactions is to better understand the organisation of the PIC. There are numerous methods to study such interactions; ELISA is only one of them. If *in vitro* translated proteins shall be used to study such interactions; the proteins should either be purified or cotranslated. Cotranslation of RT, IN and others of the PIC-members could give interesting answers.

The dynamics of the PIC can also be studied with the Biacore system. This system enables investigations of macromolecular organisations and transient interactions. Further developments of the yeast two-hybrid system are also a good means of studying the PIC.

References

1. **Pavlakis, G.**, "The Molecular Biology of Human Immunodeficiency Virus Type 1", in *AIDS: biology, diagnosis, treatment and prevention*, J. DeVita , S. Hellman, and S. Rosenberg, Editors. 1996, Lippincott-Raven: Philadelphia. p. 45-74

2. **Nermut, M.V., *et al.***, "Further evidence of icosahedral symmetry in human and simian immunodeficiency virus". *AIDS Res Hum Retroviruses*, 1993. **9**(10): p. 929-38.

3. **Marx, P.A., R.J. Munn, and K.I. Joy**, "Computer emulation of thin section electron microscopy predicts an envelope-associated icosadeltahedral capsid for human immunodeficiency virus". *Lab Invest*, 1988. **58**(1): p. 112-8.

4. **Ozel, M., G. Pauli, and H.R. Gelderblom**, "The organization of the envelope projections on the surface of HIV". *Arch Virol*, 1988. **100**(3-4): p. 255-66.

5. **Luciw, P.A.**, "Human Immunodeficiency Viruses and their replication", in *Fields virology*, B.N. Fields, *et al.*, Editors. 1996, Raven Press.: New York. p. 1881-1952.

6. **Berger, E.A., P.M. Murphy, and J.M. Farber**, "Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease". *Annu Rev Immunol*, 1999. **17**: p. 657-700.

7. **Sattentau, Q.J. and J.P. Moore**, "Conformational changes induced in the human immunodeficiency virus envelope glycoprotein by soluble CD4 binding". *J Exp Med*, 1991. **174**(2): p. 407-15.

8. **Sattentau, Q.J., *et al.***, "Conformational changes induced in the envelope glycoproteins of the human and simian immunodeficiency viruses by soluble receptor binding". *J Virol*, 1993. **67**(12): p. 7383-93.

9. **Sattentau, Q.J., S. Zolla-Pazner, and P. Poignard**, "Epitope exposure on functional, oligomeric HIV-1 gp41 molecules". *Virology*, 1995. **206**(1): p. 713-7.

10. **Karageorgos, L., P. Li, and C. Burrell**, "Characterization of HIV replication complexes early after cell-to-cell infection". *AIDS Res Hum Retroviruses*, 1993. **9**(9): p. 817-23.

11. **Bukrinsky, M.I., *et al.***, "Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection". *Proc Natl Acad Sci U S A*, 1993. **90**(13): p. 6125-9.

12. **Farnet, C.M. and W.A. Haseltine**, "Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex". *J Virol*, 1991. **65**(4): p. 1910-5.

13. **Miller, M.D., C.M. Farnet, and F.D. Bushman**, "Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition". *J Virol*, 1997. **71**(7): p. 5382-90.

14. **Pauza, C.D.**, "Two bases are deleted from the termini of HIV-1 linear DNA during integrative recombination". *Virology*, 1990. **179**(2): p. 886-9.

15. **Pomerantz, R.J., O. Bagasra, and D. Baltimore**, "Cellular latency of human immunodeficiency virus type 1". *Curr Opin Immunol*, 1992. **4**(4): p. 475-80.

16. **Cullen, B.R. and W.C. Greene**, "Regulatory pathways governing HIV-1 replication". *Cell*, 1989. **58**(3): p. 423-6.

17. **Trono, D., M.B. Feinberg, and D. Baltimore**, "HIV-1 Gag mutants can dominantly interfere with the replication of the wild-type virus". *Cell*, 1989. **59**(1): p. 113-20.

18. **Chazal, N., *et al.***, "Human immunodeficiency virus type 1 MA deletion mutants expressed in baculovirus-infected cells: cis and trans effects on the Gag precursor assembly pathway". *J Virol*, 1995. **69**(1): p. 365-75.

19. **South, T.L., *et al.***, "The nucleocapsid protein isolated from HIV-1 particles binds zinc and forms retroviral-type zinc fingers". *Biochemistry*, 1990. **29**(34): p. 7786-9.

20. **Huang, Y., *et al.***, "Incorporation of excess wild-type and mutant tRNA(3Lys) into human immunodeficiency virus type 1". *J Virol*, 1994. **68**(12): p. 7676-83.

21. **Kaplan, A.H., M. Manchester, and R. Swanstrom**, "The activity of the protease of human immunodeficiency virus type 1 is initiated at the membrane of infected cells before the release of viral proteins and is required for release to occur with maximum efficiency". *J Virol*, 1994. **68**(10): p. 6782-6.

22. **Dewar, R.L., *et al.***, "Biosynthesis and processing of human immunodeficiency virus type 1 envelope glycoproteins: effects of monensin on glycosylation and transport". *J Virol*, 1989. **63**(6): p. 2452-6.

23. **Earl, P.L., R.W. Doms, and B. Moss**, "Oligomeric structure of the human immunodeficiency virus type 1 envelope glycoprotein". *Proc Natl Acad Sci U S A*, 1990. **87**(2): p. 648-52.

24. **Dubay, J.W., *et al.***, "Analysis of the cleavage site of the human immunodeficiency virus type 1 glycoprotein: requirement of precursor cleavage for glycoprotein incorporation". *J Virol*, 1995. **69**(8): p. 4675-82.

25. **Willey, R.L., *et al.***, "Mutations within the human immunodeficiency virus type 1 gp160 envelope glycoprotein alter its intracellular transport and processing". *Virology*, 1991. **184**(1): p. 319-29.

26. **Bowerman, B., *et al.***, "A nucleoprotein complex mediates the integration of retroviral DNA". *Genes Dev*, 1989. **3**(4): p. 469-78.

27. **Bukrinsky, M.I., *et al.***, "Active nuclear import of human immunodeficiency virus type 1 preintegration complexes". *Proc Natl Acad Sci U S A*, 1992. **89**(14): p. 6580-4.

28. **Gallay, P., *et al.***, "HIV nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase". *Cell*, 1995. **83**(4): p. 569-76.

29. **Gallay, P., *et al.***, "HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway". *Proc Natl Acad Sci U S A*, 1997. **94**(18): p. 9825-30.

30. **Heinzinger, N.K., *et al.***, "The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells". *Proc Natl Acad Sci U S A*, 1994. **91**(15): p. 7311-5.

31. **Gallay, P., *et al.***, "Role of the karyopherin pathway in human immunodeficiency virus type 1 nuclear import". *J Virol*, 1996. **70**(2): p. 1027-32.

32. **Farnet, C.M. and F.D. Bushman**, "HIV-1 cDNA integration: requirement of HMG I(Y) protein for function of preintegration complexes in vitro". *Cell*, 1997. **88**(4): p. 483-92.

33. **Roe, T., *et al.***, "Integration of murine leukemia virus DNA depends on mitosis". *Embo J*, 1993. **12**(5): p. 2099-108.

34. **Lewis, P.F. and M. Emerman**, "Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus". *J Virol*, 1994. **68**(1): p. 510-6.

35. **Emerman, M., M. Bukrinsky, and M. Stevenson**, "HIV-1 infection of non-dividing cells - reply". *Nature*, 1994. **369**(6476): p. 108.

36. **Haffar, O.K., *et al.***, "CNI-H0294, a nuclear importation inhibitor of the human immunodeficiency virus type 1 genome, abrogates virus replication in infected activated peripheral blood mononuclear cells". *Antimicrob Agents Chemother*, 1998. **42**(5): p. 1133-8.

37. **von Schwedler, U., R.S. Kornbluth, and D. Trono**, "The nuclear localization signal of the matrix protein of human immunodeficiency virus type 1 allows the establishment of infection in macrophages and quiescent T lymphocytes". *Proc Natl Acad Sci U S A*, 1994. **91**(15): p. 6992-6.

38. **Popov, S., *et al.***, "Viral protein R regulates nuclear import of the HIV-1 pre-integration complex". *Embo J*, 1998. **17**(4): p. 909-17.

39. **Fouchier, R.A., *et al.***, "Interaction of the human immunodeficiency virus type 1 Vpr protein with the nuclear pore complex". *J Virol*, 1998. **72**(7): p. 6004-13.

40. **Popov, S., *et al.***, "Viral protein R regulates docking of the HIV-1 preintegration complex to the nuclear pore complex". *J Biol Chem*, 1998. **273**(21): p. 13347-52.

41. **Vodicka, M.A., *et al.***, "HIV-1 Vpr interacts with the nuclear transport pathway to promote macrophage infection". *Genes Dev*, 1998. **12**(2): p. 175-85.

42. **Bukrinsky, M.I., *et al.***, "A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells [see comments]". *Nature*, 1993. **365**(6447): p. 666-9.

43. **Fouchier, R.A., *et al.***, "HIV-1 infection of non-dividing cells: evidence that the amino-terminal basic region of the viral matrix protein is important for Gag processing but not for post-entry nuclear import". *Embo J*, 1997. **16**(15): p. 4531-9.

44. **Freed, E.O., G. Englund, and M.A. Martin**, "Role of the basic domain of human immunodeficiency virus type 1 matrix in macrophage infection". *J Virol*, 1995. **69**(6): p. 3949-54.

45. **Reil, H., *et al.***, "Efficient HIV-1 replication can occur in the absence of the viral matrix protein". *Embo J*, 1998. **17**(9): p. 2699-708.

46. **Haffar, O.K., *et al.***, "Two nuclear localization signals in the HIV-1 matrix protein regulate nuclear import of the HIV-1 pre-integration complex". *J Mol Biol*, 2000. **299**(2): p. 359-68.

47. **Le Grice, S.**, "Human immunodeficiency virus reverse transcriptase", in *Reverse transcriptase*, A. Skalka and S. Goff, Editors. 1993, Cold Spring Harbour Laboratory Press: Cold Spring Harbour, NY. p. 163-191.

48. **Kohlstaedt, L.A., *et al.***, "Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor". *Science*, 1992. **256**(5065): p. 1783-90.

49. **Hostomsky, Z., *et al.***, "Reconstitution in vitro of RNase H activity by using purified N-terminal and C-terminal domains of human immunodeficiency virus type 1 reverse transcriptase". *Proc Natl Acad Sci U S A*, 1991. **88**(4): p. 1148-52.

50. **Le Grice, S.F., *et al.***, "Subunit-selective mutagenesis indicates minimal polymerase activity in heterodimer-associated p51 HIV-1 reverse transcriptase". *Embo J*, 1991. **10**(12): p. 3905-11.

51. **Arnold, E., *et al.***, "Structure of HIV-1 reverse transcriptase/DNA complex at 7 A resolution showing active site locations". *Nature*, 1992. **357**(6373): p. 85-9.

52. **Oyama, F., *et al.***, "Intrinsic properties of reverse transcriptase in reverse transcription. Associated RNase H is essentially regarded as an endonuclease". *J Biol Chem*, 1989. **264**(31): p. 18808-17.

53. **Jacobo-Molina, A., *et al.***, "Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA". *Proc Natl Acad Sci U S A*, 1993. **90**(13): p. 6320-4.

54. **Johnson, M.S., *et al.***, "Computer analysis of retroviral pol genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes". *Proc Natl Acad Sci U S A*, 1986. **83**(20): p. 7648-52.

55. **Burke, C.J., *et al.***, "Structural implications of spectroscopic characterization of a putative zinc finger peptide from HIV-1 integrase". *J Biol Chem*, 1992. **267**(14): p. 9639-44.

56. **Zheng, R., T.M. Jenkins, and R. Craigie**, "Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity". *Proc Natl Acad Sci U S A*, 1996. **93**(24): p. 13659-64.

57. **Lee, S.P., *et al.***, "Zn2+ promotes the self-association of human immunodeficiency virus type-1 integrase in vitro". *Biochemistry*, 1997. **36**(1): p. 173-80.

58. **Heuer, T.S. and P.O. Brown**, "Mapping features of HIV-1 integrase near selected sites on viral and target DNA molecules in an active enzyme-DNA complex by photo-cross-linking". *Biochemistry*, 1997. **36**(35): p. 10655-65.

59. **Vincent, K.A., *et al.***, "Characterization of human immunodeficiency virus type 1 integrase expressed in Escherichia coli and analysis of variants with amino-terminal mutations". *J Virol*, 1993. **67**(1): p. 425-37.

60. **Asante-Appiah, E. and A.M. Skalka**, "HIV-1 integrase: structural organization, conformational changes, and catalysis". *Adv Virus Res*, 1999. **52**: p. 351-69.

61. **Cai, M., *et al.***, "Solution structure of the N-terminal zinc binding domain of HIV-1 integrase [published erratum appears in Nat Struct Biol 1997 Oct;4(10):839-40]". *Nat Struct Biol*, 1997. **4**(7): p. 567-77.

62. **Engelman, A. and R. Craigie**, "Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro". *J Virol*, 1992. **66**(11): p. 6361-9.

63. **Kulkosky, J., *et al.***, "Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases". *Mol Cell Biol*, 1992. **12**(5): p. 2331-8.

64. **Leavitt, A.D., L. Shiue, and H.E. Varmus**, "Site-directed mutagenesis of HIV-1 integrase demonstrates differential effects on integrase functions in vitro". *J Biol Chem*, 1993. **268**(3): p. 2113-9.

65. **Dyda, F., *et al.***, "Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases [see comments]". *Science*, 1994. **266**(5193): p. 1981-6.

66. **Eijkelenboom, A.P., *et al.***, "The DNA-binding domain of HIV-1 integrase has an SH3-like fold". *Nat Struct Biol*, 1995. **2**(9): p. 807-10.

67. **Woerner, A.M. and C.J. Marcus-Sekura**, "Characterization of a DNA binding domain in the C-terminus of HIV-1 integrase by deletion mutagenesis". *Nucleic Acids Res*, 1993. **21**(15): p. 3507-11.

68. **Vink, C., A.M. Oude Groeneger, and R.H. Plasterk**, "Identification of the catalytic and DNA-binding region of the human immunodeficiency virus type I integrase protein". *Nucleic Acids Res*, 1993. **21**(6): p. 1419-25.

69. **Lutzke, R.A., C. Vink, and R.H. Plasterk**, "Characterization of the minimal DNA-binding domain of the HIV integrase protein". *Nucleic Acids Res*, 1994. **22**(20): p. 4125-31.

70. **Khan, E., *et al.***, "Retroviral integrase domains: DNA binding and the recognition of LTR sequences [published erratum appears in Nucleic Acids Res 1991 Mar 25;19(6):1358]". *Nucleic Acids Res*, 1991. **19**(4): p. 851-60.

71. **Engelman, A., A.B. Hickman, and R. Craigie**, "The core and carboxyl-terminal domains of the integrase protein of human immunodeficiency virus type 1 each contribute to nonspecific DNA binding". *J Virol*, 1994. **68**(9): p. 5911-7.

72. **Lodi, P.J., *et al.***, "Solution structure of the DNA binding domain of HIV-1 integrase". *Biochemistry*, 1995. **34**(31): p. 9826-33.

73. **Chow, S.A., *et al.***, "Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus". *Science*, 1992. **255**(5045): p. 723-6.

74. **van Gent, D.C., *et al.***, "DNA binding properties of the integrase proteins of human immunodeficiency viruses types 1 and 2". *Nucleic Acids Res*, 1991. **19**(14): p. 3821-7.

75. **Wu, X., *et al.***, "Human immunodeficiency virus type 1 integrase protein promotes reverse transcription through specific interactions with the nucleoprotein reverse transcription complex". *J Virol*, 1999. **73**(3): p. 2126-35.

76. **Trentin, B., N. Rebeyrotte, and R.Z. Mamoun**, "Human T-cell leukemia virus type 1 reverse transcriptase (RT) originates from the pro and pol open reading frames and requires the presence of RT-RNase H (RH) and RT-RH-integrase proteins for its activity". *J Virol*, 1998. **72**(8): p. 6504-10.

77. **Katz, R.A. and A.M. Skalka**, "The retroviral enzymes". *Annu Rev Biochem*, 1994. **63**: p. 133-73.

78. **Davies, J.F.d., *et al.***, "Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase [see comments]". *Science*, 1991. **252**(5002): p. 88-95.

79. **Hu, S.C., *et al.***, "Murine leukemia virus pol gene products: analysis with antisera generated against reverse transcriptase and endonuclease fusion proteins expressed in Escherichia coli". *J Virol*, 1986. **60**(1): p. 267-74.

80. **Fields, S. and O. Song**, "A novel genetic system to detect protein-protein interactions". *Nature*, 1989. **340**(6230): p. 245-6.

81. **Chien, C.T., *et al.***, "The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest". *Proc Natl Acad Sci U S A*, 1991. **88**(21): p. 9578-82.

82. **Furnes, B.**, "Application of the yeast two-hybrid system to study HIV-1 integrase and reverse transcriptase interactions". 1998, Bergen.: [B.Furnes].

83. **Szilvay, A.M., *et al.***, "Epitope mapping of HIV-1 reverse transcriptase with monoclonal antibodies that inhibit polymerase and RNase H activities". *J Acquir Immune Defic Syndr*, 1992. **5**(7): p. 647-57.

84. **Nilsen, B.M., *et al.***, "Monoclonal antibodies against human immunodeficiency virus type 1 integrase: epitope mapping and differential effects on integrase activities in vitro". *J Virol*, 1996. **70**(3): p. 1580-7.

85. **Greenway, A., *et al.***, "Human immunodeficiency virus type 1 Nef binds directly to Lck and mitogen-activated protein kinase, inhibiting kinase activity". *J Virol*, 1996. **70**(10): p. 6701-8.

86. **Beckler, G.S. and R. Hurst**, "tRNAnscendTM Non-radioactive detection of in vitro translation products labeled using biotinylated lysine tRNA". *Promega Notes Magazine*, 1993(43): p. 25-32.

87. **Steinberg, T.H., *et al.***, "SYPRO orange and SYPRO red protein gel stains: one-step fluorescent staining of denaturing gels for detection of nanogram levels of protein". *Anal Biochem*, 1996. **239**(2): p. 223-37.