Mutation of Methionine 178 to Isoleucine of HIV-1 Integrase for analyzing its interaction with LEDGF/p75

By

SHAZIA SIDDIQUI
(BO12M1006)

A Dissertation Submitted to
Indian Institute of Technology Hyderabad
In Partial Fulfillment of the Requirements for
The Degree of Master of Technology

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad
Department of Biotechnology

June, 2014

Declaration

I declare that this written submission represents my ideas in my own words, and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources that have thus not been properly cited, or from whom proper permission has not been taken when needed.

SHAZIA SIDDIQUI
BO12M1006

Approval Sheet

This thesis entitled **"Mutation of Methionine 178 to Isoleucine of HIV-1 Integrase for analyzing its interaction with LEDGF/p75"** by Shazia Siddiqui is approved for the degree of Master of Technology from IIT Hyderabad.

Dr. Shashidhar
Assistant Professor
Department of Civil Engineering
IIT Hyderabad

Dr. Anindya Roy
Assistant Professor
Department of Biotechnology
IIT Hyderabad

Dr. N.K. Raghavendra
Assistant Professor
Department of Biotechnology
IIT Hyderabad

Dr. Rajakumara Eerappa
Assistant Professor
Department of Biotechnology
IIT Hyderabad

Acknowledgements

With deep regards and profound respect, I would like to thank my supervisor Dr. N.K Raghavendra for his encouragement, enthusiasm, support and guidance during my thesis work. It has been a pleasure and a true privilege to work with him. I would also like to thank my thesis committee: Dr. Anindya Roy, Dr. Basant Kumar Patel for their insightful comments, and valuable suggestions. I would also like to take this opportunity to thank all members of NKR laboratory: Anu George, Manne Aravind Kumar, Venkateswara Rao Nuthakki and Sree Hari who have always been there when I needed help with my experiments, enquiries; providing motivation and a delightful working environment. Also thanks to my friends for their moral support. In the last, I would like to thank my parents, aunty and my sisters, for their moral and financial support, and their belief in me.

The study was supported by grants from Department of Biotechnology, New Delhi, to Dr. N.K.Raghavendra.

Abstract

HIV-1 replication depends on the insertion of a cDNA copy of the viral genome into the host genome. This process is catalyzed by the viral integrase (IN), in a two-step process-

3'processing and strand transfer reaction. Lens epithelium-derived growth factor (LEDGF)/p75 is an important cellular co-factor for human immunodeficiency virus-1 (HIV-1) replication which plays a crucial role during HIV-1 cDNA integration. The C-terminal integrase binding domain (IBD) of LEDGF/p75 interacts with CCD of HIV-1 Integrase (IN). In this study, based on the co-crystal structure of IN-LEDGF/p75, Methionine residue at position 178 of HIV-1 IN has been mutated to isoleucine. The M178I mutant IN awas analyzed for interaction with IBD of LEDGF/p75 in the pull down assays. Compared to wild type IN, M178I mutant show a ~90% reduction in interaction with IBD. Pull down result shows that M178I is an important residue which participates in interaction with IBD of LEDGF/p75.

Abbreviations

IN	Integrase
NTD	N-terminal domain
CCD	catalytic core domain
CTD	C-terminal domain
HT	His tag
PIC	preintegration complex
IBD	Integrase binding domain
LEDGF/p75	Lens epithelium derived growth factor
HIV-1	Human Immunodeficiency Syndrome
MA	matrix protein
CA	Capsid protein
NC	nucleocapsid protein
SP1	spacer peptide 1,
SP2	spacer peptide 2
PR	protease
RT	reverse transcriptase

Contents

Chapter 1

INTRODUCTION

HIV-1 is a Lentivirus, a member of Retrovirus family which causes the acquired immunodeficiency syndrdrome (AIDS)(Weiss 1993). HIV infects immune cells in the human immune system such as helper T cells (CD4+ T cells), macrophages, and dendritic cells. These cells act as reservoir of HIV virus and increase the latency period and lead to AIDS (Stevenson 2003). Infection of HIV-1 virus causes decrease in the CD4+ cells. When $CD4^+$ T cell numbers decline below a critical level, cell mediated immunity is lost, and the body becomes progressively more susceptible to infections

1.1 STRUCTURE OF HIV-1 VIRUS

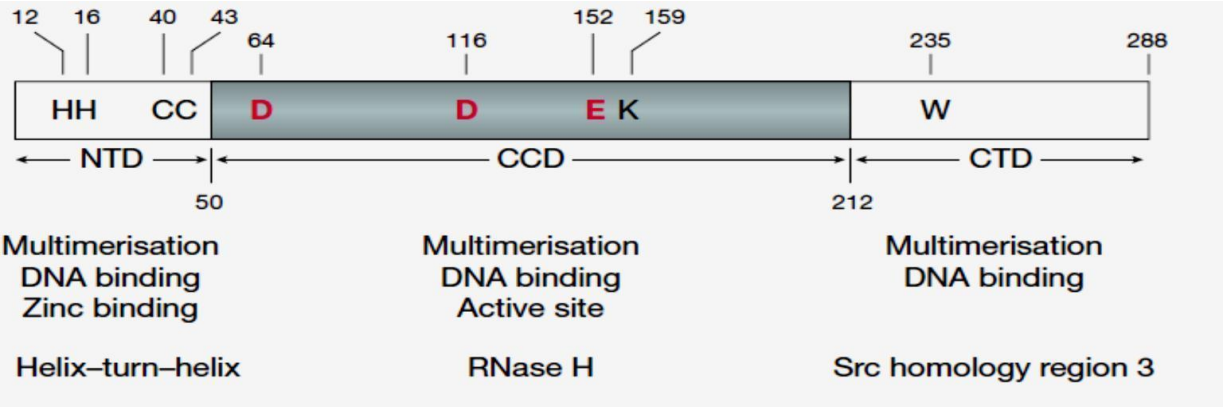
HIV-1 is roughly spherical. The outermost covering is a lipid bilayer called the envelope, which is derived from the plasma membrane of the human host cell where the virus was produced. The outer surface of the HIV-1 contains a type of surface proteins composed of the two peptides gp41 and gp120. Portion of gp120 interacts with the CD4 receptor on the human host cell and is the first step in the life cycle of HIV-1. Below the lipid envelope of HIV-1 there are two different layers of protein. The outer protein layer is known as the matrix and is made up of the MA protein. The inner protein layer is called the capsid and is composed of the CA protein. The capsid surrounds the core, which, in turn, contains two copies of the viral RNA genome and associated viral proteins including reverse transcriptase and integrase.

HIV-1 contains two copies of the viral RNA genome which encodes nine genes (gag, pol, tat, env, rev, nef, vif, vpr and vpu). These nine genes encode 15 proteins. Gag encodes Gag polyprotein. The polyprotein is subsequently proteolysed by viral protease into MA (matrix protein, p17), CA (capsid protein, p24), NC (nucleocapsid protein, p7) and p6. Pol encodes three proteins PR (protease), IN (integrase) and RT (reverse transcriptase). Gene env encodes two proteins, SU and TM, these two proteins are structural component and form outer membrane envelope. Virus encodes additional proteins called accessory proteins: vif, vpr, vpu and Nef present in viral particle. Tat and Rev work as regulatory proteins (Frankel and Young 1998).

1.2 HIV-1 INTEGRASE

HIV-1Integrase IN is a 32 kDa protein with 288 amino acids encoded by *pol* gene as a part of Gag-Pol polyprotein which is processed by HIV-1 protease (Delelis et al. 2008). Integrase belongs to superfamily of polynucleotidyl transferases which includes other proteins like Mu transposase and RNaseH (Dyda et al. 1994). Integrase enzyme catalyzes the integration of the reverse transcribed viral DNA into the host genome (Bushman and Craigie 1991, LaFemina, et al. 1991, Engelman and Craigie 1992), which is essential step of virus

replication. IN has a three domain structure: N-terminal domain (NTD), Catalytic core domain (CCD) and C-terminal domain (CTD) (Figure1) (Engelman and Craigie 1992, Engelman et al. 1993, van Gent et al. 1993).



Fiqure1: Domain organisation and amino acid residues conserved among HIV-1 integrase proteins. The enzyme comprises the N-terminal domain (NTD), residues 1–49, catalytic core domain (CCD), residues 50–212 and C-terminal domain (CTD) residues 213– 288. The aspartic acid (D64, D116) and glutamic acid (E152) residues in the CCD form the DDE motif (red font), play a major role in integration. Lysine (K) 159 a conserved residue in transposase protein, binds to the viral LTR. Tryptophan (W) 235 in the CTD is conserved only among the retroviral integrases. (Adapted from (Vandegraaff and Engelman, 2007)

N-terminal domain (residue 1-51) is formed by a bundle of three α-helices and has a HHCC motif (Cai, Zheng et al. 1997), which is conserved in retroviral integrases. The HHCC motif indicates Histidine and Cysteine residues H12, H16, C40 and C43 which interacts with zinc with a stoichiometry of 1:1 and stabilizes the interaction between the helices (Zheng, Jenkins et al. 1996). HHCC motif is proposed as zinc finger structure based on the similarity to the CCHH motif found in transcription factor TFIIIA. Zinc binding to HHCC motif promotes multimerization and enhances the catalytic activity of IN (Zheng, Jenkins et al. 1996). Substitution of any one of the conserved histidine or cysteine residues in this motif would be expected to disrupt the zinc finger, but mutation of IN at these residues does not completely abolish catalytic activity *in vitro* (Cannon et al. 1994).

Catalytic Core domain (residue 52-211) has DD35E motif (figure 2) which is conserved among retroviral integrase, and transposase protein. These acidic residues (D64, D116, E152) co-ordinate with Mg^{2+} or Mn^{2+} (Goldgur, Dyda et al. 1998) and catalyse the 3` processing and strand transfer activities of viral DNA. Mutants of active site residues (D64, D52, E116) show most drastic effect on all *in vitro* IN activities such as 3` processing, strand transfer and disintegration. (Engelman & Craigie, 1992; Kulkosky et al. 1992, van Gent et al. 1993). CCD also takes part in multimerization, CCD-CCD interface is fairly large (~1,650 Å2) and include Glu85, Ala86, Tyr99, Phe100, Lys103, Trp108, Gln177 and Asn184 residues at the interface (Christ et al. 2010).

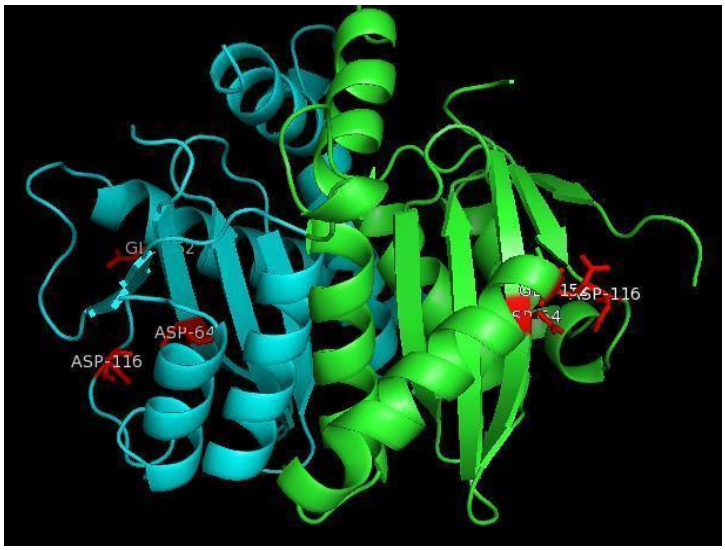


Figure 2. Structure of catalytic core dimer (Pdb Id 1BIS): CCD is formed by five-stranded mixed β sheet flanked by helices. The catalytic residues of the D, D (35) E sequence couloured in red and these residues are located in opposite face of the CCD dimerization interface. Dimerization of the CCD domain involves predominantly hydrophobic interactions. (Mariusz et al 2011)

C-terminal domain (residue 212-288) is least conserved domain. NMR structure of CTD is dimeric with each monomer consisting of five β strands (residue 222-229, 232-245, 248-253, 256-260, 265-268) arranged in an antiparallel manner (figure 3). Topology of CTD structure resembles the Src homology3 (SH3) domain (Eijkelenboom et al. 1995) with highly basic amino acid residues interacting with DNA in a nonspecific manner (Chen et al. 2000).

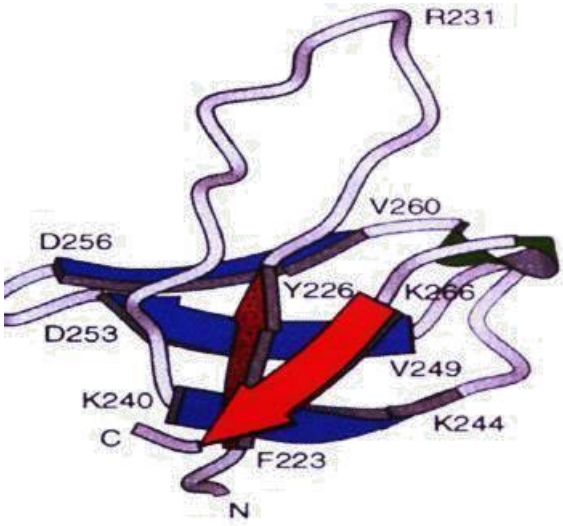


Figure 3. Monomeric structure of C-terminal Domain. Three stranded β sheet coloured in blue and two stranded β sheet coloured in red are shown which are connected by the loops. Length of strand is denoted by the labelled residue. (Adapted from (Eijkelenboom, Lutzke et al. 1995)

1.3 TWO–DOMAIN STRUCTURE OF HIV-1 INTEGRASE

Two-domain crystal structures comprised of the CCD and CTD (Chen et al 2000) or NTD and CCD (Wang et al. 2001) showed dimeric organizations. In two domain structure of

NTD-CCD, NTD is linked to the CCD through a linker segment of amino acids 47–55 shown in Figure 4 (right) with an arrow (Wang et al. 2001). Another flexible loop comprised of residues 187–194 connects the CCD with the CTD is shown in Figure 4 (left) (Chen et al. 2000). Dimer interfaces for individual NTDs and CTDs with CCD differs for two domain system (Kessl et al. 2009). Dimer can catalyse *in vitro* 3` processing and strand transfer reaction but fails in concerted integration because the active sites of integrase are located on opposite sides of the dimeric CCD structure separated by 40 A °, while the distance between target scissile bonds in ideal B form DNA is close to 18 A ° so tetramer of a integrase (Hare et al. 2009) (figure 5) which is a compact structure is necessary to accomplish concerted integration of both LTRs (8 –10 bp) into host DNA. IN tetramer is a dimer-of-dimer and tetramerisation is primarily mediated by intermolecular interactions between NTD-CCD (Wang et al. 2001, Hare et al. 2009). Within tetramer a pair of inner NTDs mediates stable dimer-dimer interactions whereas remaining (outer) NTDs do not participate in the tetramerisation. The NTD-CCD interface within the monomer includes a well-defined salt bridge between Glu11 and Lys186 and hydrophobic interactions involving Tyr15, Trp132, Val165, Phe181, and Ile182. And an additional salt bridge is formed between Asp25 and Lys188 shown in figure 6 (Hare et al. 2009).

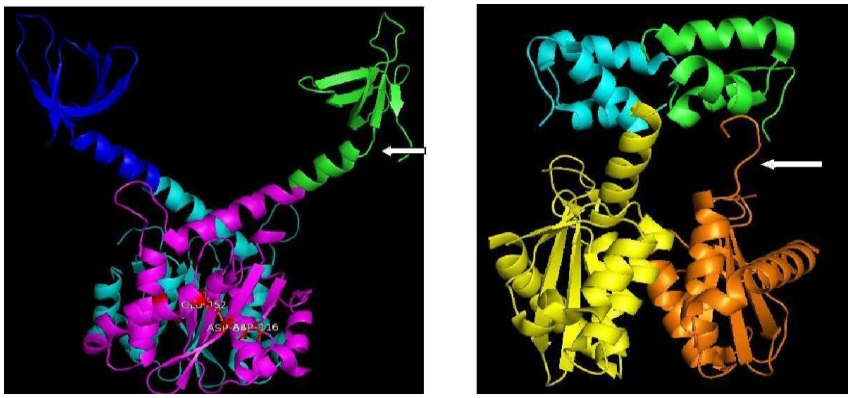
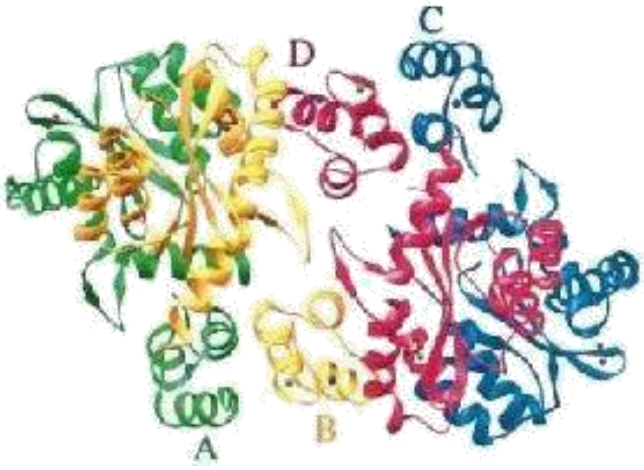


Figure 4. Two–domain structure of HIV-1 integrase. Left figure. HIV-1 CCD-CTD (pdb Id 1EX4) dimer formed by two monomers (green+purple and blue+cyan). catalytic core domain for two monomers are shown in purple and cyan colour dimer and C-terminal domain are shown in green and blue. Right figure. HIV-1(1-212) dimer (pdb Id K6Y) formed by two monomer (green+orange and blue+yellow). NTD are coloured blue and green) and CCD are coloured yellow and orange. (Adpted from (Chen, Krucinski et al. 2000, Wang, Ling et al. 2001)



Fiqure 5. Ribbons diagram of the integrase tetramer. The integrase full-length tetramer formed by superposition of the catalytic core domain of the HIV-1 IN52-288 structure (PDB code 1EX4) onto the catalytic domains of IN1-212. A,B,C and D reprasents the four monomers having CCD and NTD are shown. (Adapted from Wang et al. 2001)

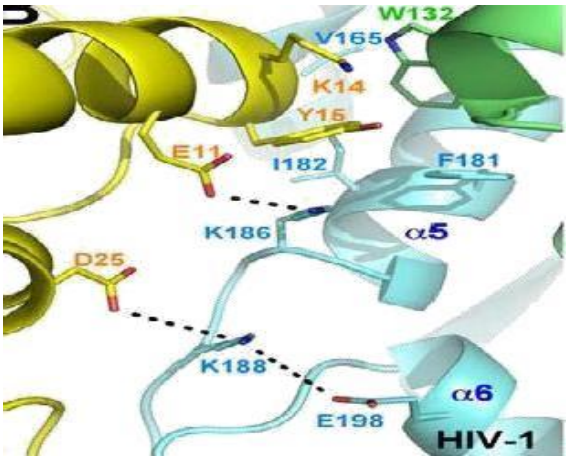


Figure 6. The NTD-CCD interfaces. NTD (yellow) of one dimer interact with CCD (blue) of other dimer. CCD (green) also takes part in interaction. Salt bridge formed between D25:E198 and E11:K186 other residue shown takes part in hydrophobic interaction. (Hare et al. 2009)

1.4 FUNCTION OF CATALYTIC CENTRE MOTIF OF INTEGRASE PROTEIN (DD35E).

DD35E Motif plays central role for integration reaction by coordinating with two divalent cations (Mg2+ or Mn2+). One metal ion functions as a general base. It increases the partial negative charge of the incoming nucleophile i.e water by deprotonation. The second metal ion functions as a general acid that assists the leaving of 5` phosphate oxygen anion of GT dinucleotide and stabilizes a

pentavalent transition state by coordination of the non-bridging oxygen atoms at the cleavage site (Figure 7) (Haren et al. 1999).

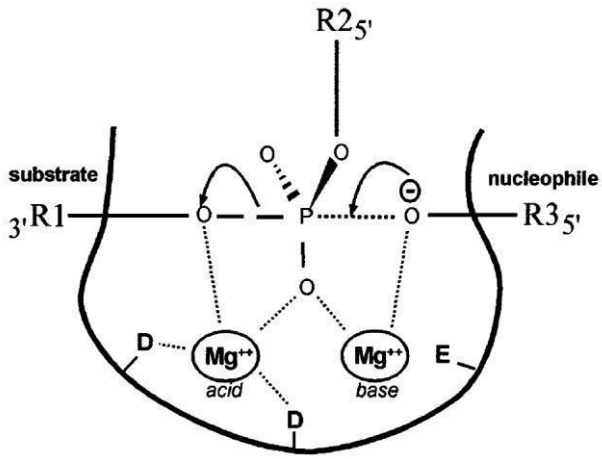


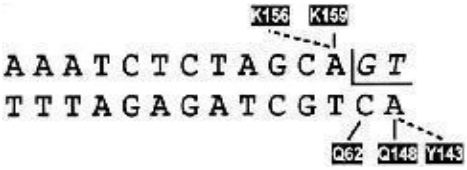
Figure 7. Role of divalent metal ions in catalysis. Catalytic pocket containing the penta coordinated phosphate in the transition state is shown. One Mg2+ ions proposed to act as a Lewis acid is coordinated with two aspartate residues (D64, D116) , the oxyanion of the leaving group and a nonbridging oxygen atom. The second Mg2+ is also shown to complex with the nonbridging oxygen and to act as a base for deprotonization of the incoming nucleophile i.e water. (Adapted from (Haren et al. 1999)

1.5 MECHANISM OF RETROVIRAL DNA INTEGRATION

Integration is a two steps process

1) 3` end processing
2) strand transfer reaction

IN binds to the short viral DNA sequence (12-20 bases) called attachment sites shown in figure 8 present at end of long terminal repeats (LTR) which initiates the 3' processing (Katzman and Katz 1999).



Fiqure 8. The HIV-1 DNA end. GT dinucleotide which undergoes 3'processing shown in italics at 3` end. Amino acids residues Q62, Q143, Y148, K156 and K159 within catalytic core domain cross links to viral DNA end and positions the viral DNA for proper 3' processing. (Adapted from Haren et al. 1999)

3` processing takes place in the cytoplasm after the cDNA is formed form viral RNA. IN uses water molecule as nucleophile to hydrolyze GT dinucleotide from each 3' end of the viral DNA and forms 3'OH which act as nucleophile in strand transfer reaction (figure 9).

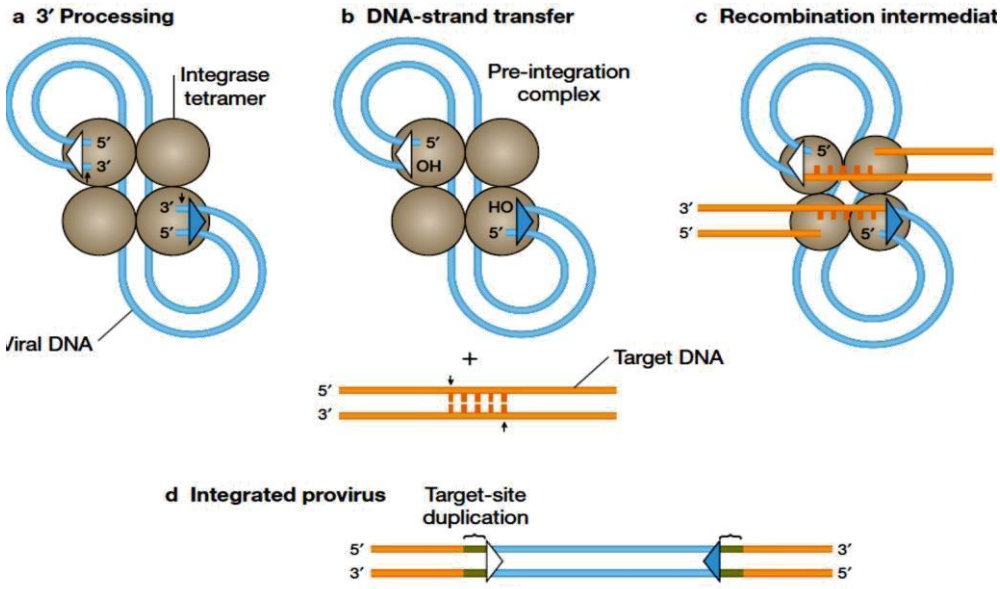


Figure 9. Mechanism of HIV-1 DNA integration. (a) A tetramer of integrase engage two ends of HIV-1 DNA. (b) In strand transfer reaction integrase uses the 3'-hydroxyl groups of the cleaved viral DNA to cut the target in a staggered fashion (vertical arrows), which joins the viral 3' ends to the 5'-phosphates of the cut (c) The recombination intermediate formed by integrase DNA-strand-transfer activity comprises joined viral 3' ends but free 5' ends. (d) Gap repair of the DNA recombination intermediate yields the integrated provirus flanked by a 5 bp duplication of target DNA (Adapted from Vandegraaff and Engelman 2007).

Strand transfer reaction takes place in nucleus and after nuclear entry of processed viral DNA, IN uses the 3` hydroxyl groups of the cleaved viral DNA to cut the target DNA at the transcriptional active sites and joins the viral 3` ends to the 5` phosphates of the cut. Integration process is completed by cleavage of unpaired dinucleotide from 5` end of viral DNA which is mediated by integrase and repair of single stranded gap with help of host repair proteins such as DNA-PK, DNA ligase IV, Ku and XRCC4.

Recombinant integrase is capable of catalysing two reactions (3` processing and strand transfer reaction) of integration process in absence of other viral and cellular proteins. However studies show that *in vivo* integration process involves cellular protein like Importin 7, Lens Epithelium Derived Growth Factor (LEDGF/p75), Barrier to Auto integration Factor (BAF), Gemin 2, High Mobility Group chromosomal protein A1 (HMGA) and viral protein like matrix, reverse transcriptase, and nucleocapsid. These viral and cellular protein form nucleoprotein

complex called preintegration complex (PIC), assisting the viral genome into cellular genome. The IN can reverse the strand transfer reaction by site selective cleaving of the integrated DNA. This reaction is called disintegration which is just reverse of strand transfer reaction occurring only *in vitro*. (Chow et al. 1992).

1.6 INTEGRASE MUTANTS

Reverse transcription products of HIV-1 RNA are linear and circular DNA. These linear DNA are substrate of viral integration and circular DNA are dead end products. Replication-defective integrase mutants of HIV-1 can be grouped into one of two phenotypic classes (Engelman 1999)

Class I Mutants specifically blocks at integration step as a result they are defective in 3'processing and strand transfer reaction. Reverse transcription is not affected and number of LTR circle is increased which acts as marker in class I mutants. Example: Mutant of catalytic triad (D64, D116, E52) prior to integration (reverse transcription) which results in the reduction in the level of viral DNA synthesis in infected cells. In addition to defective reverse transcription, HIV-1 particle assembly and release from infected cells are also blocked. Example: H12L, H12N mutations.

1.7 LEDGF

LEDGF/p75 is a member of the hepatoma-derived growth factor (HDGF) related protein (HRP) family ((Dietz et al. 2002). HRPs are characterized by a conserved N-terminal PWWP domain (Pro-Trp-Trp-Pro) (Stec et al. 2000,Qiu et al. 2002). Six human HRP family members have been described: HDGF, HRP1, HRP2, HRP3, LEDGF/p75, and LEDGF/p52 (Izumoto et al. 1997, Ikegame et al. 1999, Cherepanov et al. 2004). LEDGF/p75 and HRP2 possess affinity for HIV-1 IN with their conserved IN-binding domain (IBD) (Cherepanov et al. 2004). LEDGF/p75 encoded by the PSIP1 gene and 530-aa protein, PSIP1 also encodes a 333-aa splice variant LEDGF/p52 by alternate splicing (Ge Si and Roeder 1998).

LEDGF/p52 and p75 share NH2-terminal amino acids 1 to 325. LEDGF/ p52 has a unique intron at COOH-terminal tail (amino acids 326-333). The NH2 terminus of both p52 and p75 contains a PWWP domain (amino acids 1-93), three charged domains, a nuclear localization signal and two AT-hook sequences (Gly-Arg-Lys-Arg-Lys-Ala-Glu-Lys-Gln) (Vanegas et al. 2005, Maertens et al. 2004) (figure 10). Binding of LEDGF/p75 to DNA in vitro is mediated by the NLS and dual copy of the AT-hook DNA binding motif, whereas the N-terminal PWWP domain supplies a critical chromatin recognition function (Turlure et al. 2006). Charged regions (CRs) 1–3 work with the PWWP domain and AT-hooks to affect chromatin binding.

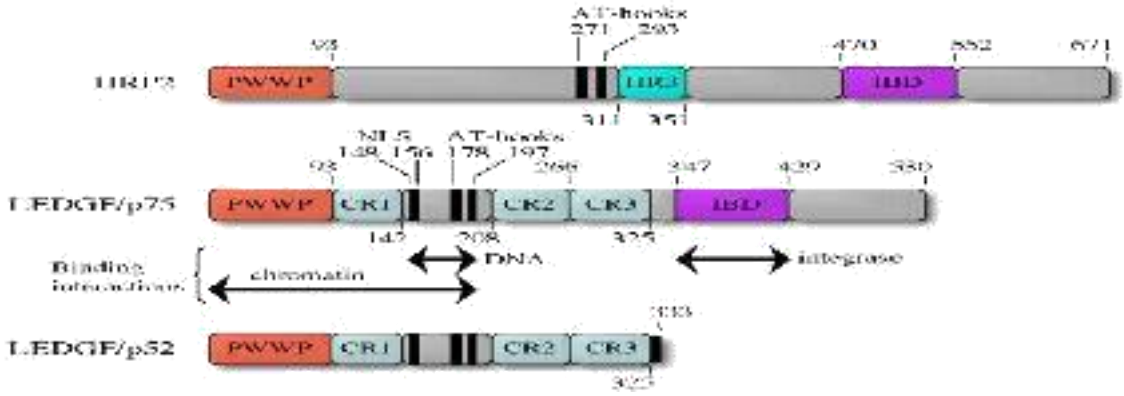


Figure 10: Domain Organization of LEDGF/p75 and related proteins (HRP2 and LEDGF/p52): The N-terminal PWWP domain is conserved supplies a critical chromatin recognition function. HRP2 and LEDGF/p75 both have IBD domain which interact with CCD domain of HIV-1 integrase. HRP2 has homology region III (HR3). LEDGF/p75 and LEDGF/p52 both have NLS, AT-hook, and three charged region. The binding of LEDGF/p75 to DNA is mediated by the NLS and a nearby dual copy of the AT-hook DNA binding motif .Charged regions (CRs) 1–3 work in concert with the PWWP domain.(Adapted from Engelman and Cherepanov 2008).

LEDGF/p75 is closely associated with condensed chromatin in the nucleus. LEDGF/p75 is involved in transcriptional activation of stress-related or anti-apoptotic proteins, such as heat shock protein 27, αB-crystallin, heat-shock protein 90 and antioxidant protein 2 (Singh, and Fatma 2002) and promotes cell growth and protects cells against stress-induced cell death. LEDGF/p75 is identified as a common nuclear auto antigen in a variety of inflammatory conditions, such as atopic dermatitis, asthma and interstitial cystitis. LEDGF/p75 protein functions as a nuclear auto antigen generating an autoantibody response in prostate cancer (Ganapathy and Casiano 2003, Daniels et al. 2005). LEDGF/p75 can be cleaved by caspases-3 and -7, which are the main effector caspases in apoptosis, resulting in two cleavage fragments of 65 and 58 kDa. These LEDGF/p75 fragments disturbed the pro-survival role of LEDGF/p75 (Wu et al. 2002).

NMR structure of the LEDGF/p75 IBD revealed a compact α-helical domain possessing topological and structural similarities to HEAT repeat domains (Cherepanov et al. 2005). HEAT repeat is an α-helical hairpin structure containing 37–47 amino acid residues, and found in diverse protein families Huntingtin, elongation factor 3, regulatory subunit of protein phosphatase 2A, and PI3-kinase TOR. LEDGF/p75 IBD is comprised of only two repeats and therefore classified as a pseudo-HEAT repeat analogous topology (PHAT) domain. IN dimer forms a cleft like structure, IBD burrows into this cleft (Figure 11). I-365, D-366, and P-409 are important residues of LEDGF which are interact with Integrase (Cherepanov et al. 2005). The side chain of LEDGF residue Ile-365 projects into a hydrophobic pocket formed by IN B-chain residues Leu-102, Ala-128, Ala-129, and Trp-132 and A-

chain residues Thr-174 and Met-178. LEDGF residues Phe-406 and Val-408 interacts Trp-131 in the IN B chain. LEDGF Asp-366 makes a bidentate hydrogen bond to the main-chain amides of IN residues Glu-170 and His-171 in chain A. The backbone amide of LEDGF residue Ile-365 makes a hydrogen bond to the backbone carbonyl group of Gln-168 in chain A (Cherepanov et al. 2005).

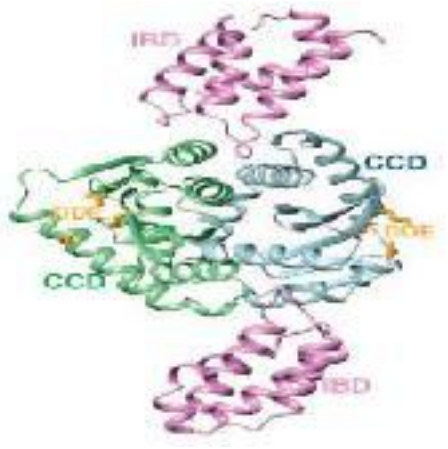


Figure 11: The overall structure of the CCD–IBD complex. IN chains A and B are colored blue and green, respectively; the IBD subunits are violet. The side chains of the DDE catalytic triad are shown as yellow sticks

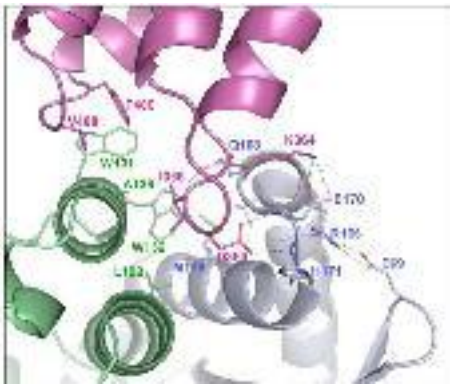


Figure 12 Crystal Structure of the LEDGF/p75-IN Interaction**:** Residues Ile-365 and Asp-366 of loop between IBD helices 1 and 2, project into a pocket at the CCD dimer interface. The bidentate hydrogen bond contact between Asp-366 and the backbone amides of IN residues Glu-170 and His-171 is critical for the protein–protein interaction in vitro, during HIV-1 infection. Ile-365 is buried into a hydrophobic pocket predominantly formed by IN residues Ala-128, Trp-132, Leu-102, and Met-178. (Adapted from Engelman and Cherepanov 2008).

LEDGF/p75 is an essential HIV-1 IN cellular cofactor for viral replication (Emiliani 2005). LEDGF acts as a chromatin tethering factor to facilitate the interaction between IN and nuclear chromatin (Maertens et al. 2003). LEDGF/p75 also directs viral DNA integration at AT-rich DNA transcriptional units of the cellular genome (Ciuffi et al. 2005, Shun et al. 2007).

The versatile functions of LEDGF/p75 in HIV 1 replication include nuclear import of IN, tethering of IN to chromatin for directing viral DNA integration into transcriptionally active regions and protection from proteosomal degradation (Maertens et al. 2003, Llano et al. 2004).

1.8 HYPOTHESIS

Human transcriptional co-activator LEDGF/p75 is an essential cellular binding partner of HIV-1 integrase. Crystal structure of IN-IBD shows that the side chain of LEDGF residue Ile-365 projects into a hydrophobic pocket formed by IN B-chain residues Leu-102, Ala-128, Ala-129, and Trp-132 and A-chain residues Thr-174 and Met- 178(Cherepanov et al. 2005). Recent study with binding energy calculation is showed that residue Met178 of IN interacts with the LEDGF/p75 integrase binding domain (IBD) (Xue 2014). Residue M178 located at the dimeric interface of IN and participate in π electron orbital interaction between four amino acids (M178, W132, F181 and F185). Sulfhydryl–aromatic distance between M178–W132 is 4.13 Å (shortest distance between S atom and ring center) (Al-Mawsawi et al. 2008)

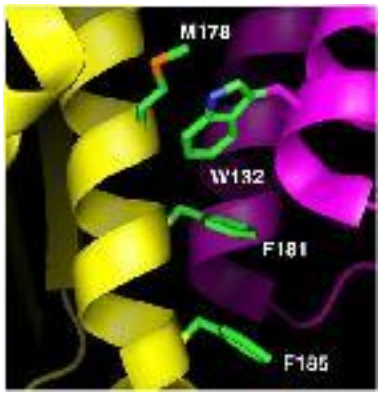


Figure13: The four-tiered aromatic interaction at the IN dimeric interface: Aromatic interaction between M178, W132, F181, and F185 at the dimeric interface (PDB 2B4J) (Adapted from Al-Mawsawi et al. 2008).

Substitution of M178 to C showed decreases in both 3` processing and strand transfer activities. M178C was replication incompetent and showed undetectable levels of 2-LTR circles and integrated proviral DNA in comparison to WT virus (Al- Mawsawi et al. 2008). In another study M178A virus was shown to be replication incompetent (Priet et al. 2003).

Two recently discovered LEDGINs BI-1001 and CX14442 (Xue 2014) and six Chemical compounds (Reddy et al 2014) are also found to interact with residue M178. Interaction of M178 residue of IN with LEDGF is only shown by computational study. In our study we substitute M178 to Isoleucine and check mutation effect on interaction with LEDGF with the help of PULL DOWN.

1.9 OBJECTIVES

1. To perform site directed mutagenesis of IN Met178 residue to Ile for analyzing its interaction with IBD of LEDGF/p75.

2. Expression and purification of mutant His-Tag Integrase mutant M178I protein in bacterial expression system.

3. To determine the interaction of the mutant IN with IBD of LEDGF/p75 using pull down assays.

Chapter 2

MATERIALS AND METHODS

2.1 CONSTRUCTION OF SITE-DIRECTED IN M178I MUTANT.

Site-directed mutagenesis was done by overlapping PCR using a two-step procedure as described below:

Plasmid DNA encoding the mutations M178I within full-length integrase was prepared using pfu DNA polymerase. In first step of PCR two sets of primers are used in two separate amplification reaction to amplify overlapping DNA fragenment.

NR1 (5` CTG GGA TCC TTT TTA GAT GGA ATA GAT AAG GCC 3`) & NR41 (CTGCTATTTGTACTGCTGTTTTAAG) primers were used in 1st amplification reaction and NR2 (5`ACG ACT CGA GCT AAT CCT CAT CCT GTC TAC TTG 3`) & NR40 (NR40:CTTAAAACAGCAGTACAAATAGCAG) primers were used in 2nd amplification reaction. Primers NR40 & NR41 have desired sequence for M178I mutation and also carry change to delete the AflII restriction site. Plasmid pINKB6HThr was used as template for both amplification reactions. These two sets of reactions produced half fragments containing the mutation site as overlap. In second step of PCR, two overlapping half PCR fragments were used as template and NR1 and NR2 used as forward and reverse primer. Reaction mixture of all three PCR reactions contains 10X Pfu buffer,10mM dNTP, Pfu Polymerase (0.5µl), 5µM of primers, 1X BSA . Conditions for all three PCR reaction were same and as follows: 1 cycle of denaturation at 95ºC for 3min, 25 cycles of denaturation at 95 ºC for 15sec, annealing at 50ºC for 15sec ,extension at 68ºC for 60sec and 1 cycle of final extension at 68ºC for 5min.

2.2 DIGESTION.

After overlap PCR, product and pFT1 LEDGF vector were digested with XhoI and BamHI in presence of buffer 4. Digestion reaction was incubated at 37ºC for three hours.

2.3 LIGATION.

Ligation of Digested product of mutant and pFT1 LEDGF vector backbone were carried out in presence of 1X ligase buffer, 10mM ATP and 0.5µl. This reaction mixture was incubated at 18ºC for 4hr.

2.4 SCREENING OF TRANSFORMANTS.

The ligated product is transformed into DH5α competent cells as follows:

1 After Adding 20µl of ligated plasmid DNA to the competent cells and gently tap the cells incubate on ice for 15 mins.

2 A heat shock was given to cells at 42°c for 90 sec

3 After heat shock cells were kept on ice for 15 mins.

4 Later 15min 1ml LB broth was added to cells and incubated for 1 hr at 37° in incubator.

5 Culture was spun for 2 mins.

6 Cells were resuspended in 100 ul media and spread onto LB plate containing Amp.

7 Plate was incubated at 37°c for overnight in incubator.

2.5 MINI PREP BY ALKALINE LYSIS METHOD.

1 DH5α colony was picked from LB plate containing Ampicillin antibiotic and inoculate into 3 ml LB broth containing Ampicillin.

2 Culture was incubate at 37°c for overnight in shaker incubator (250 rpm)

3 Next day cells were harvested from 1.5 ml culture at 14.8 k for 2 mins at room temperature.

4 pellet was resuspended in 100µl of ice cold alkaline lysis solution I (25mM Glucose, 10mM EDTA, 50Mm Tris-chloride pH =8)

5 200 µl of freshly prepared alkaline lysis solution II (0.2 N NaOH, 1%SDS) was added and invert gently for 5 times and kept on ice for 5mins.

6 150 µl of ice cold Alkaline lysis solution III (it is an acidic solution with pH 4.8 (potassium acetate) was added and invert gently, kept on ice for 5 mins.

7 After Centrifugation at 14.8 k for 5 mins at 4°c supernatant was transferred into the fresh tube.

8 Equal volumes of phenol and chloroform in 1:1 ratio (i.e 225ul of phenol and 225 ul of chloroform) was added to the supernatant and after vortexing centrifuged at 14.8 k for 5 mins at room temperature.

9 Aqueous upper layer was transferred to fresh tube and add 0.1 volumes of 3M Sodium acetate (i.e 45ul) and 2 volumes of absolute ethanol mixed it by vortexing.

10 After centrifugation at 14.8 k for 10 mins at 4°c pellet was washed with 70% Ethanol.

11 Pellet was kept for drying at 37°c incubator for 15 min

15 Dried pellet was dissolved in 50µl of TE buffer (10mMTris cl PH 8,1mM EDTApH8) and added 1µl of RNase.

2.6 SCREENING OF CLONE

Screening of clone was done by PCR DNA isolated from mini prep method was used as template DNA. Reaction mixture contains 1X Pfu buffer,1mM dNTP, Pfu Pol, 0.2 µM of each NR1 and NR2 primers and 1X BSA. Reaction condition was similar as for overlap PCR. PCR product was checked on 0.6% agrose gel.

2.7 SCREENING OF MUTANT

pRA12 mutant was screen by AflII digestion. Two reactions were set, one as control with pINKB6HThr (wild type) and other with pRA12 (mutant type). Control reaction has 2µg pINKB6HThr and mutant has 2µg of DNA isolated from mini prep method (pRA12). Reaction product was incubated at 37°c incubator for 5 hr. Digestion product was checked on 0.6% agrose gel at 50Volts.

2.8 INSERT RELEASE

For insert release two reactions were set, one for wild type pINKB6H and other for mutant pRA12. Digestion of pRAKB6HThr was performed with XhoI and XbaI. Digestion of pRA12 was done with XhoI and BamHI.

2.9 EXPRESSION OF PROTEIN

For expression of mutant IN, DNA isolated from Midi prep method was transformed in BL21 DE3 pLyS competent cells. Transformed colony was inoculated in 3ml LB broth containing 50µg /ml ampicillin (3 µl of 50mg/ml stock) and 30µg/ml chloramphenicol (3µl of 30mg/ml) and incubated overnight at 37°c 250 rpm. Next day 5% preinoculam was used for expression of mutant protein in 200ml LB broth containing 50µg/ml ampicillin. And induction was done by IPTG (1 mM) at absorbance of 0.6–0.7 optical density at 600 nm. After induction, culture was allowed to grow for an additional 4 h at 28 °C. This was followed by centrifugation of the cells, at 5000 rpm at 24°C for 15min. Cell pellet was resuspended in 5ml lysis buffer (50mM TrisCl p7.4, 15mM imidazole, 1MNaCl, 7.5mM CHAPS, 0.5mM PMSF). The resuspended pellet was incubated for 30 min at room temperature after adding 0.3mg/ml lysozyme. The cells were lysed with sonication. Lysate was centrifuged at 14.8k, 4°C for 15min. The supernatant was loaded onto a 1 ml Ni2+-nitrilotriacetic acid agarose column pre-equilibrated with buffer containing 15mM imidazole, 1M NaCl and 50mM TrisCl. After 1hour nutation of beads with supernatant, beads were loaded onto column, column was washed with 30 column volumes of wash buffer containing 50mM imidazole, 1M NaCl and 50mM TrisCl. Integrase was eluted with elution buffer containing 400mM imidazole, 1MNaCl, 7.5mM CHAPS and 0.5mM PMSF. Elution fraction having mutant IN were dialyzed against buffer 500 mM NaCl, 50 mM Tris-Cl pH 7.5. Dialyzed mutant protein was stored in storage buffer having 0.5mM PMSF, 1X cocktail inhibitor, 5mM DTT, 7.5mM CHAPS, 50%glycerol at -20°C. Protein concentration was determined by the Bradford method with purified BSA as a standard.

PULL DOWN

Binding of IN to IBD was assayed in pull-down buffer (PB): 150 mM NaCl, 25 mM Imidazole, 2mM $MgCl_2$, 0.1% NP40, 50 mM Tris HCl pH 7.4. A 4 µg of recombinant His_6-tagged HIV-1 IN were incubated with 3 µg of IBD in 200 µl of PB and followed by 30min nutation at 18°C. Samples were supplemented with 50 µl (settled beads volume) of Ni-charged resin.The mixtures were kept for nutation at 18°C. Beads were recovered by centrifugation and washed with 1 ml of ice-cold PB. Beads were boiled in 12µl 1XTris-glycineand 5µl dye. Eluted proteins were separated by SDS-PAGE.

Chapter 3

RESULT AND DISCUSSION

3.1 SITE DIRECTED MUTAGENESIS OF HIV-1 IN MET178 RESIDUE TO ILE:

IN1-288/M178I mutant was generated by overlap PCR. In first step of overlap PCR two fragments were generated using two sets of primers forward/antisense and sense/reverse primers of which sense and antisense will carry nucleotide change for M178I. In the second step of overlap PCR two fragments obtained in first step of overlap PCR were used as templates with NR1—NR2 primer sets to generate full length IN M178I. PCR products were checked on agarose gel, shown in figure 14. DNA fragment produced with NR1/NR41 primers showed band near 500bp and fragment produced with NR2/NR40 primers showed band near 300bp shown in Figure 14A. Overlap PCR product with NR1-NR2 showed band between 800 and 900bp as shown in Fig 14 B.

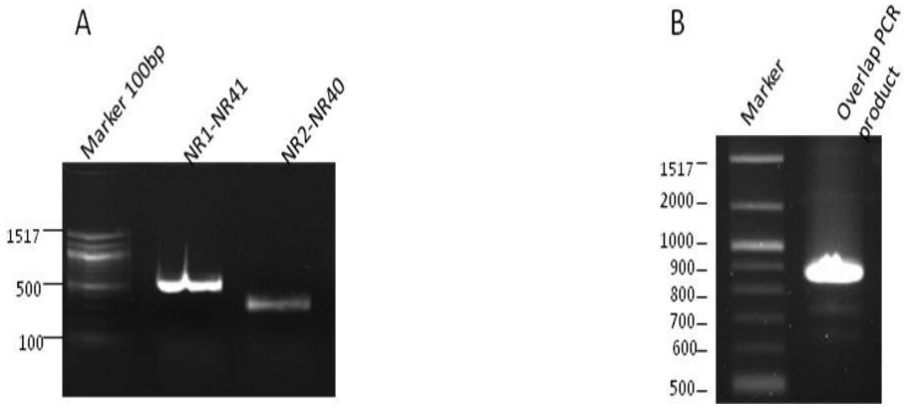


Figure 14: Overlap PCR for generating IN M178I mutation: A) First step of overlap PCR using NR1-NR41 and NR2-NR42 generating two fragments around 500bp and 300bp. B) Second step of overlap PCR using NR1-NR2 generating full length Mutant IN M178I of expected size 864bp.

3.2 SCREENING OF CLONE AND MUTATION.

Eight colonies were obtained after transformation of ligated product of pFT1 LEDGF and mutant IN M178I. The colonies were picked into 3ml of LB with

ampicillin overnight. Mini prep plasmid DNA isolation was done with 1.5ml culture. After isolation, plasmid DNA was checked on agarose gel as shown in Fig 15. Colony 4 and 6 DNA showed lower mobility as compared to pFT1 LEDGF/p75. These were screened further by PCR amplification using NR1-NR2 primer set. PCR product was checked on agrose gel which showed that clone 4 and clone 6 formed expected band of 864bp length as shown in Fig 16.

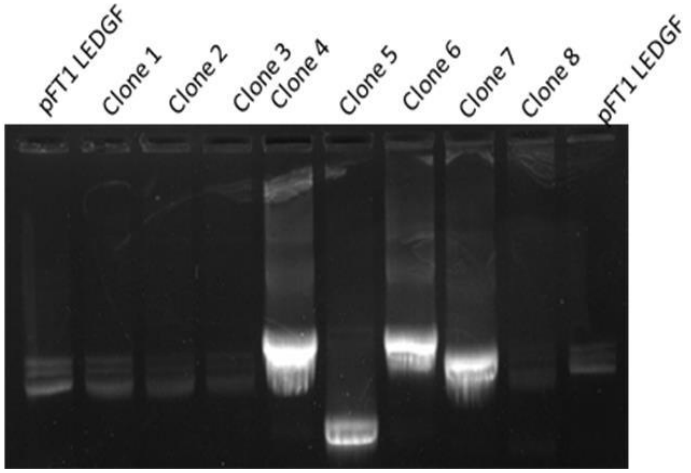


Figure 15 Mini prep plasmid DNA isolation. DNA isolated from eight colonies were checked on gel. pFT1 LEDGF was used as Control. Clone 1, 2 3, 7, and 8 were showing same mobility as that of pFT1 LEDGF. Clone 4 and Clone 6 showed lower mobility as compared with pFT1 LEDGF.

Insert release was checked by digestion of pRA12 with XhoI and BamHI and for digestion of pINKB6HThr was with XhoI and XbaI enzymes. For pRA12 digestion with XhoI and BamHI an insert release of 864 bp and for pINKB6HThr digestion with XhoI and XbaI an insert release of 915bp was obtained (fig 17).

For screening M178I mutant in IN, AflII digestion was performed, pRA12 (plasmid from colony 4) loses AflII site after generating M178I mutation and wild type IN (pINKB6HThr) retains AflII site. After digestion with AflII, a linear fragment was obtained for pINKB6HThr while pRA12 showed undigested nicked and supercoiled DNA, as can be seen in Fig 18 confirming the presence of M178I mutation in colony 4 DNA of pRA12.

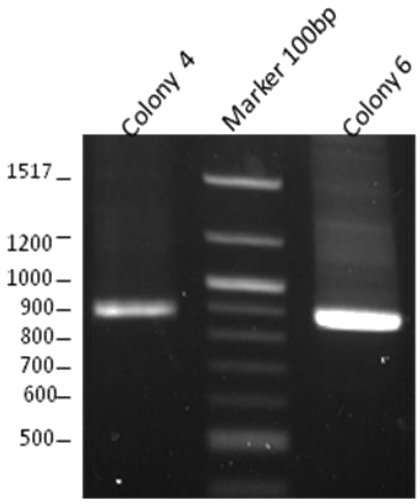


Figure 16 Screening of Clone using PCR: NR1-NR2 PCR amplification product of clones 4 and 6 mini DNA was shown. Both the clones showed the expected amplification product of 864 bp. The 100bp ladder used as marker.

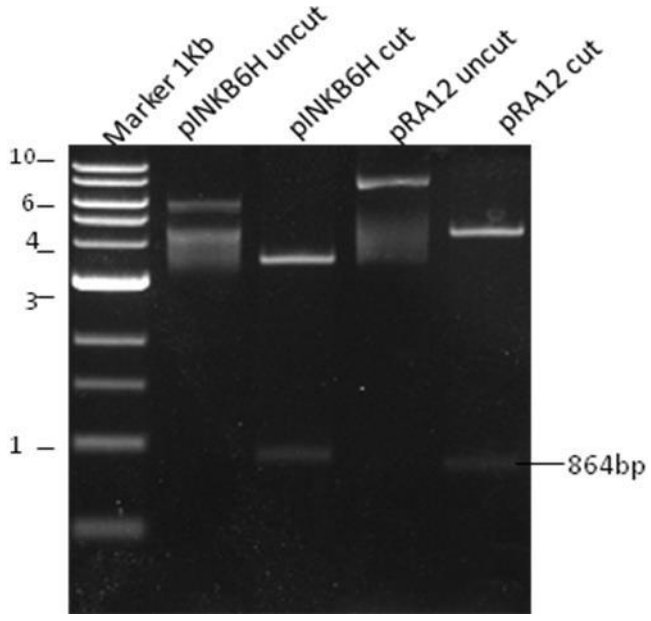


Figure 17 Insert release of pRA12. Insert release was analyzed on agrose gel with 1Kb ladder as marker. pINKB6HThr uncut is used as control . pINKB6HThr digested with XhoI and XbaI showed insert release of 915 bp and pRA12 digested with XhoI and BamHI showed the insert release of size 864 bp, as expected.

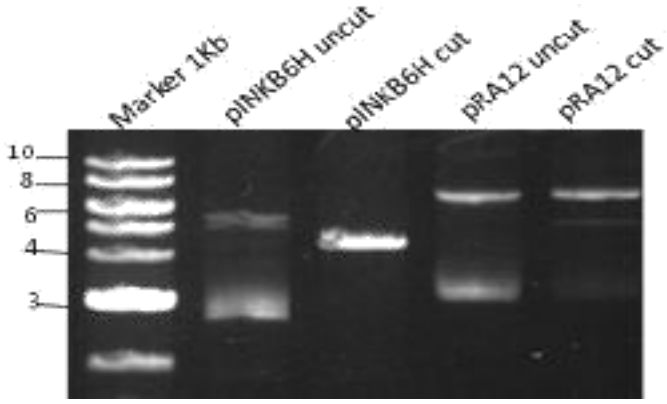


Figure 18 AflII digestion of pRA12**.** Digestion of pRA12 was analysed on agrose gel with 1 Kb marker. pINKB6HThr with AflII site digested with AflII restriction enzyme gave linearized vector where as pINKB6HThr uncut showed two forms supercoiled and nicked one on the gel. pRA12 cut with AflII restriction enzyme shown same band pattern as uncut pRA12, which indicates the loss of AflII site as a result of generating M178I mutation in IN.

3.3 PROTEIN PURIFICATION OF IN M178I:

IN1-288/M178I mutant protein was purified from the soluble fraction after lysis of cells. His Tag at N-terminus of the IN M178I protein permitted rapid purification by nickel-affinity chromatography as described in materials and methods, homogenous preparation of M178I mutant of IN was obtained. Purification was analyzed by SDS-PAGE (fig 19).

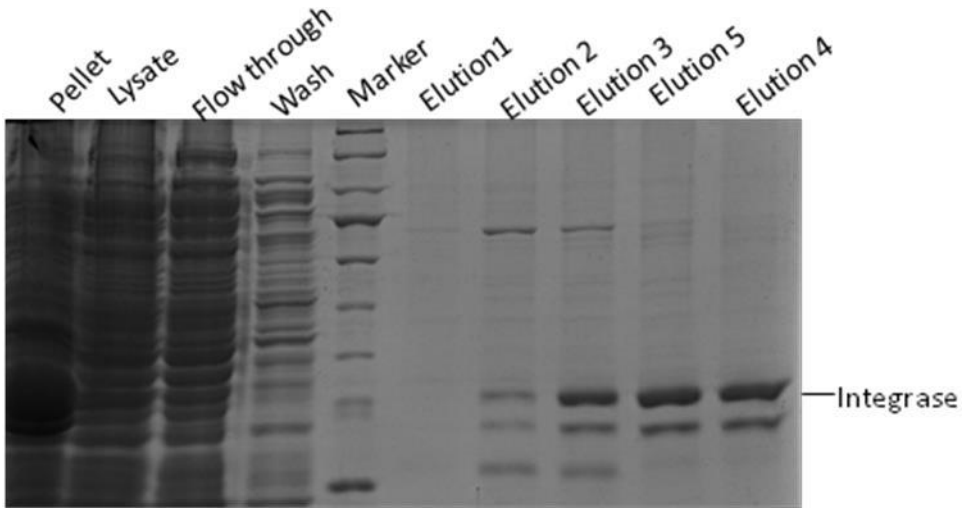


Figure 19 Protein purifation of pRA12. Protein is eluted with 400mM imidazole and fractions were analyzed by SDS-PAGE. Pellet, Lysate, flow through, wash and 5 elution fractions are loaded on gel. Protein ladder was used as Marker. Protein was dialyzed from elution fraction 4 and 5 and used for assays.

3.4 Pull Down

Pull down of IBD using His6-tag IN was done for analysis of interaction. Wild type IN with wild type IBD and wild type IN with mutant IBD D366N were used as positive and negative controls. Wild IN shows interaction with wild type IBD but did not show interaction with D366N mutant of IBD, as was demonstrate earlier in literature (Cherepanov et al 2005). Mutant IN M178I shows negligible interaction with wild type IBD as can be seen in figure 20. Quantification of pull down assay indicates a 93% reduction in interaction between IBD of LEDGF/p75 and M178I mutant of IN as shown in graph 1.

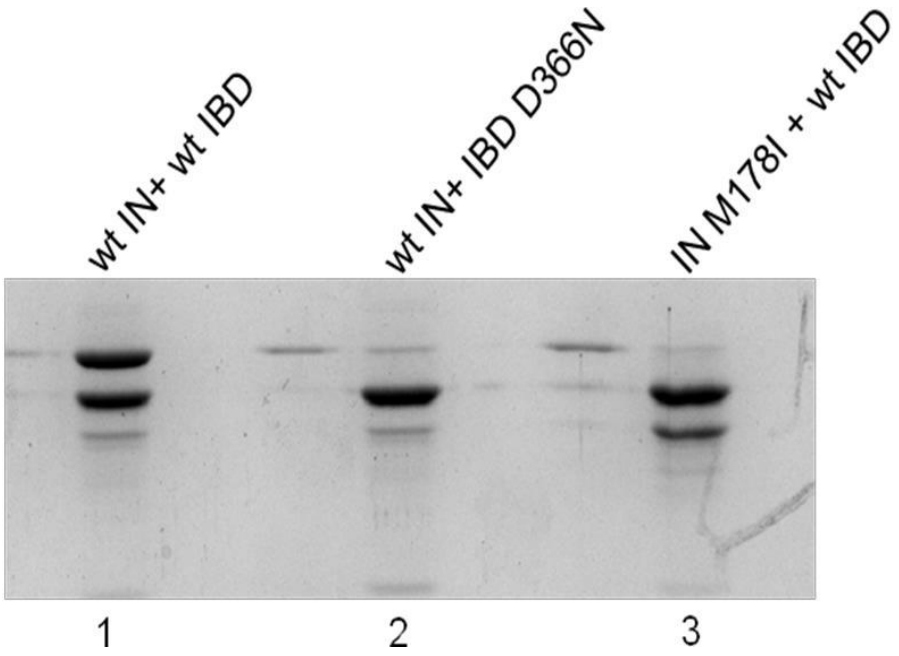
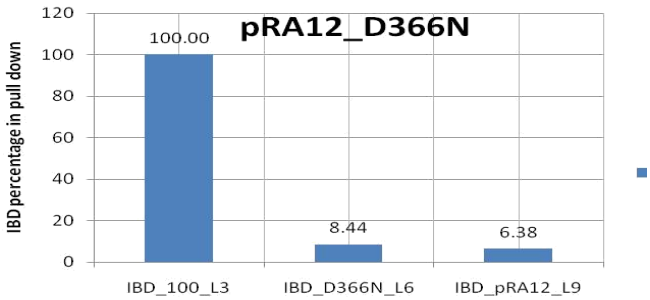


Figure 20 His6-tag pull-down assay for the N-terminally His6-tagged IN mutants. The IN mutant proteins were incubated with IBD and Ni2+-NTA agarose beads. The pulled-down proteins were checked on separated on a SDS–PAGE gel and. Wild type IN with wild type IBD and wild type IN with mutant IBD D366N were used as control. Wild IN shows interaction with type IBD but did not show interaction with mutant IBD D366N. Mutant IN M178I shows reduced interaction with wild type IBD.



Graph 1 Quantification of pull down of IBD with mutant His6 tag IN M178I. Wild type IN interaction with wild type IBD is set to 100% and with mutant IBD D366N serves as negative control. Mutant IN M178I shows a 93% loss of interaction with wild type IBD of LEDGF/p75.

3.5 DISCUSSION

LEDGF/p75 is an important cellular binding partner of HIV-1 integrase. Residue Ile 365 of LEDGF projects into a hydrophobic pocket formed by IN B-chain residues Leu-102, Ala-128, Ala-129, and Trp-132 and A-chain residues Thr-174 and Met-178. With the help of pull down assay, this study shows that M178 plays an important role in interaction between LEDGF/p75 and IN. Pull down assay shows that mutation of IN at M178 results in a nearly 90% loss of interaction with IBD.

Chapter 4

CONCLUSION

LEDGF/p75 is important cellular co factor for HIV-1 replication. LEDGF/p75 interacts with HIV-1 IN. In this study, HIV-1 IN has been mutated at residue M178 using overlap PCR and interaction of purified mutant with IBD of LEDGF/p75 was analyzed. A 90% decrease in interaction between M178I mutant of IN and wild type IBD of LEDGF/p75 indicates that M178 is makes crucial interactions with I365 of IBD of LEDGF/p75.

References

Al-Mawsawi, L. Q., F. Christ, R. Dayam, Z. Debyser, and N. Neamati. 2008. Inhibitory profile of a LEDGF/p75 peptide against HIV-1 integrase: insight into integrase-DNA complex formation and catalysis. *FEBS Lett* 582 (10): 1425-30.

Al-Mawsawi, L. Q., A. Hombrouck, R. Dayam, Z. Debyser, and N. Neamati. 2008. Four-tiered pi interaction at the dimeric interface of HIV-1 integrase critical for DNA integration and viral infectivity. *Virology* 377 (2): 355-63.

Cherepanov, P., A. L. Ambrosio, S. Rahman, T. Ellenberger, and A. Engelman. 2005. Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. *Proc Natl Acad Sci U S A* 102 (48): 17308-13.

Cherepanov, P., E. Devroe, P. A. Silver, and A. Engelman. 2004. Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. *J Biol Chem* 279 (47): 48883-92.

Cherepanov, P., Z. Y. Sun, S. Rahman, G. Maertens, G. Wagner, and A. Engelman. 2005. Solution structure of the HIV-1 integrase-binding domain in LEDGF/p75. *Nat Struct Mol Biol* 12 (6): 526-32.

Ciuffi, A., M. Llano, E. Poeschla, C. Hoffmann, J. Leipzig, P. Shinn, J. R. Ecker, and F. Bushman. 2005. A role for LEDGF/p75 in targeting HIV DNA integration. *Nat Med* 11 (12): 1287-9.

Daniels, T., J. Zhang, I. Gutierrez, M. L. Elliot, B. Yamada, M. J. Heeb, S. M. Sheets, X. Wu, and C. A. Casiano. 2005. Antinuclear autoantibodies in prostate cancer: immunity to LEDGF/p75, a survival protein highly expressed in prostate tumors and cleaved during apoptosis. *Prostate* 62 (1): 14-26.

Dietz, F., S. Franken, K. Yoshida, H. Nakamura, J. Kappler, and V. Gieselmann. 2002. The family of hepatoma-derived growth factor proteins: characterization of a new member HRP-4 and classification of its subfamilies. *Biochem J* 366 (Pt 2): 491-500.

Engelman, A., and P. Cherepanov. 2008. The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication. *PLoS Pathog* 4 (3): e1000046.

Frankel, A. D., and J. A. Young. 1998. HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem* 67: 1-25.

Ganapathy, V., T. Daniels, and C. A. Casiano. 2003. LEDGF/p75: a novel nuclear autoantigen at the crossroads of cell survival and apoptosis. *Autoimmun Rev* 2 (5): 290-7.

Ge, H., Y. Si, and R. G. Roeder. 1998. Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. *EMBO J* 17 (22): 6723-9.

Ikegame, K., M. Yamamoto, Y. Kishima, H. Enomoto, K. Yoshida, M. Suemura, T. Kishimoto, and H. Nakamura. 1999. A new member of a hepatoma-derived growth factor gene family can translocate to the nucleus. *Biochem Biophys Res Commun* 266 (1): 81-7.

Izumoto, Y., T. Kuroda, H. Harada, T. Kishimoto, and H. Nakamura. 1997. Hepatoma-derived growth factor belongs to a gene family in mice showing significant homology in the amino terminus. *Biochem Biophys Res Commun* 238 (1): 26-32.

Llano, M., S. Delgado, M. Vanegas, and E. M. Poeschla. 2004. Lens epithelium-derived growth factor/p75 prevents proteasomal degradation of HIV-1 integrase. *J Biol Chem* 279 (53): 55570-7.

Maertens, G., P. Cherepanov, Z. Debyser, Y. Engelborghs, and A. Engelman. 2004. Identification and characterization of a functional nuclear localization

signal in the HIV-1 integrase interactor LEDGF/p75. *J Biol Chem* 279 (32): 33421-9.

Maertens, G., P. Cherepanov, W. Pluymers, K. Busschots, E. De Clercq, Z. Debyser, and Y. Engelborghs. 2003. LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *J Biol Chem* 278 (35): 33528-39.

Priet, S., J. M. Navarro, G. Querat, and J. Sire. 2003. Reversion of the lethal phenotype of an HIV-1 integrase mutant virus by overexpression of the same integrase mutant protein. *J Biol Chem* 278 (23): 20724-30.

Qiu, C., K. Sawada, X. Zhang, and X. Cheng. 2002. The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. *Nat Struct Biol* 9 (3): 217-24.

Reddy, K. K., P. Singh, and S. K. Singh. 2014. Blocking the interaction between HIV-1 integrase and human LEDGF/p75: mutational studies, virtual screening and molecular dynamics simulations. *Mol Biosyst* 10 (3): 526-36.

Shinohara, T., D. P. Singh, and N. Fatma. 2002. LEDGF, a survival factor, activates stress-related genes. *Prog Retin Eye Res* 21 (3): 341-58.

Shun, M. C., N. K. Raghavendra, N. Vandegraaff, J. E. Daigle, S. Hughes, P. Kellam, P. Cherepanov, and A. Engelman. 2007. LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev* 21 (14): 1767-78.

Stec, I., S. B. Nagl, G. J. van Ommen, and J. T. den Dunnen. 2000. The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? *FEBS Lett* 473 (1): 1-5.

Stevenson, M. 2003. HIV-1 pathogenesis. *Nat Med* 9 (7): 853-60.

Turlure, F., G. Maertens, S. Rahman, P. Cherepanov, and A. Engelman. 2006. A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo. *Nucleic Acids Res* 34 (5): 1653-65.

Vanegas, M., M. Llano, S. Delgado, D. Thompson, M. Peretz, and E. Poeschla. 2005. Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering. *J Cell Sci* 118 (Pt 8): 1733-43.

Weiss, R. A. 1993. How does HIV cause AIDS? *Science* 260 (5112): 1273-9.

Wu, X., T. Daniels, C. Molinaro, M. B. Lilly, and C. A. Casiano. 2002. Caspase cleavage of the nuclear autoantigen LEDGF/p75 abrogates its pro-survival function: implications for autoimmunity in atopic disorders. *Cell Death Differ* 9 (9): 915-25.

Xue, W., H. Liu, and X. Yao. 2014. Molecular modeling study on the allosteric inhibition mechanism of HIV-1 integrase by LEDGF/p75 binding site inhibitors. *PLoS One* 9 (3): e90799.