

Alma Mater Studiorum – Università di Bologna

**DOTTORATO DI RICERCA IN
BIOLOGIA CELLULARE E MOLECOLARE**

Ciclo XXVI

Settore Concorsuale di afferenza: 05 / I1

Settore Scientifico disciplinare: BIO / 19

**COSTRUZIONE DI CHIMERE MOLECOLARI
DELL'ENZIMA INTEGRASI DI HIV CON ATTIVITA'
IDROLIZZANTE DELLE LTR VIRALI.**

Presentata da:

Dott.ssa Regina Muthusamy

Coordinatore Dottorato

Chiar.mo Prof. Vincenzo Scarlato

Relatore

**Chiar.ma Prof.ssa
Gabriella Campadelli-Fiume**

Relatore

Dott. Alessandro Ripalti

Esame finale anno 2014

INDICE

RIASSUNTO	5
------------------	---

COSTRUZIONE DI CHIMERE MOLECOLARI DELL'ENZIMA INTEGRASI DI HIV CON ATTIVITA' IDROLIZZANTE DELLE LTR VIRALI.

PARTE GENERALE

1.	INTRODUZIONE	7
1.1	AIDS	7
1.1.1	CARATTERISTICHE GENERALI	7
1.1.2	STORIA DELL'AIDS	8
1.1.3	CENNI DI EPIDEMIOLOGIA	9
1.2	HIV	10
1.2.1	SOTTOTIPI DI HIV	10
1.2.2	MORFOLOGIA DEL VIRIONE	11
1.2.3	IL GENOMA DI HIV-1	12
1.2.4	PROTEINE STRUTTURALI DI HIV-1	14
1.2.5	PROTEINE REGOLATRICI ED ACCESSORIE DI HIV-1	16
1.2.6	CICLO REPLICATIVO DI HIV-1	19
1.3	INTEGRASI	24
1.3.1	STRUTTURA	24
1.3.2	FUNZIONE DELL'INTEGRASI	27
1.3.3	SITI DI INTEGRAZIONE	29
1.3.4	COFATTORI	32
1.4	TERAPIA FARMACOLOGICA E HAART	34
1.5	APPROCCI ALTERNATIVI	36
1.5.1	GENOME EDITING	37
1.5.2	NUCLEASI ZINC FINGER (ZFN)	39
1.5.3	FOKI	42

PARTE SPERIMENTALE

2.	OBIETTIVO DELLA RICERCA	44
3.	MATERIALI E METODI	46
3.1	COLTURE BATTERICHE	46
3.1.1	CEPPI BATTERICI	46
3.1.2	TERRENI DI COLTURA	47
3.1.3	PREPARAZIONE DEI BATTERI ELETTROCOMPETENTI	48
3.1.4	TRASFORMAZIONE DEI BATTERI PER ELETTROPORAZIONE	48
3.1.5	PURIFICAZIONE DEL DNA PLASMIDICO DA <i>E.Coli</i>	49
3.2.	BIOLOGIA MOLECOLARE	50
3.2.1	PLASMIDI	50
3.2.2	OLIGONUCLEOTIDI	51
3.2.3	AMPLIFICAZIONE MEDIANTE PCR	52
3.2.4	MUTAGENESI SITO-SPECIFICA	53
3.2.4.1	MUTAGENESI INDOTTA MEDIANTE RADIAZIONI UV	55
3.2.5	PURIFICAZIONE DI FRAMMENTI DI DNA	55
3.2.6	ELETTROFORESI SU GEL DI AGAROSIO E SCREENING DEI TRASFORMANTI	56
3.2.7	LIGAZIONE	56
3.2.8	PROVE DI ESPRESSIONE	57
3.2.9	SDS PAGE	58
3.2.10	WESTERN BLOT	59
3.2.11	PURIFICAZIONE DELLA CHIMERA	60
3.2.12	EMSA	63
4.	RISULTATI	64
4.1	STRATEGIA E COSTRUZIONE DELLA CHIMERA	64
4.2	STRATEGIA E COSTRUZIONE BERSAGLI	67
4.3	ESPRESSIONE E PURIFICAZIONE DELLA CHIMERA	68
4.4	EMSA	70
4.5	SISTEMA DI SELEZIONE IN VIVO	74
5.	DISCUSSIONE	84
6.	BIBLIOGRAFIA	86

**ANALISI *IN SILICO* DI SEGNALI DI LOCALIZZAZIONE
NUCLEARE SUL PROTEOMA DI CHLAMYDIA SPP. E STUDIO
FUNZIONALE DI DUE NLS DELLA PROTEINA EFFETTRICE
PMPD**

	FAMIGLIA CHLAMYDIACEAE	93
1.	INTRODUZIONE	93
2.	MATERIALI E METODI	97
3.	RISULTATI	99
4.	CONCLUSIONI	102
5.	BIBLIOGRAFIA	104

RIASSUNTO

La Sindrome da Immunodeficienza Acquisita (AIDS o SIDA) causata da HIV-1 (Virus dell'Immunodeficienza umana) è caratterizzata dalla graduale compromissione del sistema immunitario del soggetto colpito. Le attuali terapie farmacologiche, purtroppo, non riescono a eliminare l'infezione a causa della comparsa di continui ceppi resistenti ai farmaci, e inoltre questi trattamenti non sono in grado di eliminare i reservoir virali latenti e permettere l'eradicazione definitiva del virus dall'organismo.

E' in questo ambito che si colloca il progetto a cui ho lavorato principalmente in questi anni, cioè la creazione di una strategia per eradicare il provirus di HIV integrato nel genoma della cellula ospite.

L'Integrasi di HIV-1 è un enzima che media l'integrazione del cDNA virale nel genoma della cellula ospite. La nostra idea è stata, quindi, quella di associare all'attività di legame dell'IN stessa, un'attività catalitica. A tal fine abbiamo creato una proteina chimerica costituita da un dominio DNA-binding, dato dall'Integrasi, e da un dominio con attività nucleasica fornito dall'enzima FokI.

La chimera ottenuta è stata sottoposta a mutagenesi random mediante UV, ed è stata oggetto di selezione in vivo, al fine di ottenere una chimera capace di riconoscere, specificamente le LTR di HIV-1, e idrolizzare i siti di inserzione.

Questo lavoro porterà a definire pertanto se l'IN di HIV può essere riprogrammata a catalizzare una nuova funzione mediante la sostituzione dell'attività del proprio dominio catalitico con quello di FokI.

**COSTRUZIONE DI CHIMERE MOLECOLARI DELL'ENZIMA
INTEGRASI DI HIV CON ATTIVITA' IDROLIZZANTE DELLE
LTR VIRALI.**

PARTE GENERALE

1. INTRODUZIONE

1.1 AIDS

1.1.1 CARATTERISTICHE GENERALI

La Sindrome da Immunodeficienza Acquisita (AIDS), è un insieme di manifestazioni patologiche causate dal virus dell'immunodeficienza umana di tipo 1(HIV-1).

L'infezione da HIV causa un progressivo indebolimento di tutte le difese immunitarie dell'organismo, in quanto il virus infetta in maniera specifica e distrugge i linfociti T CD4+ coinvolti nella risposta immunitaria, facilitando in tal modo l'insorgenza di gravi patologie, come cancro (frequente è il sarcoma di Kaposi), encefaliti o comparsa di malattie infettive causate da patogeni opportunisti che si sviluppano nei pazienti debilitati dalla malattia.

L'HIV si trasmette attraverso lo scambio di fluidi corporei infetti: sangue, liquido seminale, secrezioni vaginali e latte materno.

Un individuo infetto è un portatore asintomatico e può in seguito sviluppare la sindrome vera e propria, l'AIDS conclamata. La progressione dell'infezione è influenzata da fattori genetici legati al virus o all'ospite, ^[1; 2] dalle condizioni igieniche, dalle co-infezioni opportunistiche e dal particolare ceppo del virus infettante. ^[3; 4]

L'infezione da HIV attualmente deve considerarsi curabile, ma non guaribile.

Oggi si dispone di numerosi farmaci antiretrovirali che aggrediscono il virus in diverse fasi del suo ciclo replicativo inibendo gli enzimi che caratterizzano tali fasi. Questi

farmaci purtroppo non possono eliminare l'infezione, ma solo “cronicizzarla”.

L'introduzione della HAART, terapia antiretrovirale altamente attiva, ha cambiato profondamente il quadro dell'infezione nei paesi sviluppati, i benefici ottenuti sono: riduzione drastica della mortalità, riduzione drastica dei casi di AIDS, aumentata sopravvivenza delle persone con HIV.

Numerosi vaccini contro il virus HIV sono attualmente in fase di sperimentazione, sia con finalità terapeutiche che per la prevenzione. Nessuno è però ancora uscito dalle fasi di sperimentazione, né è prevedibile quando ciò possa eventualmente verificarsi.

La creazione di un vaccino risulta molto difficile a causa dell'enorme diversità genetica del virus e della difficoltà nell'indurre la formazione di anticorpi neutralizzanti, quindi le strategie più recenti prevedono piuttosto l'induzione di un'immunità in grado di controllare la progressione della malattia, allungando così i tempi di sopravvivenza e allontanando l'insorgenza dell'AIDS.

1.1.2 STORIA DELL'AIDS

Il virus dell'HIV ha avuto origine in Africa all'inizio degli anni '30 in seguito all'adattamento del Virus dell'Immunodeficienza delle Scimmie (SIV, Simian Immunodeficiency Virus) alla specie umana. ^[5] Grazie agli studi di biologia molecolare si è potuto stabilire una relazione fra l'HIV ed il SIV, identificando una omologia genetica del 98% tra questi due virus, ed arrivando a costruire un vero e proprio albero genealogico. La trasmissione del virus dalla scimmia all'uomo sarebbe avvenuta in seguito ai contatti tra le due specie attraverso modalità rese possibili dallo stretto contatto che l'uomo aveva con il sangue di tali animali (pratiche alimentari, sacrificali ecc). L'infezione si sarebbe poi diffusa al continente europeo ed agli Stati Uniti soprattutto per mezzo dei flussi migratori.

I primi casi di AIDS si osservarono nel 1981 negli Stati Uniti d'America quando tra

giovani omosessuali si verificarono diversi casi di polmoniti da *Pneumocystis Carinii* e sarcoma di Kaposi, patologie normalmente molto rare nella popolazione sana. L'anno seguente, dato l'incremento del numero di questi casi, le autorità sanitarie americane coniarono il termine di AIDS (*Acquired ImmunoDeficiency Syndrome*) per questa nuova patologia. L'agente eziologico HIV è stato identificato e isolato nel 1983 quasi contemporaneamente dal gruppo di Luc Montagnier dell'Istituto Pasteur di Parigi e da Robert Gallo del National Cancer Institute di Bethesda. Nel 1985 la FDA (Food and Drug Administration) approvò il primo test per la determinazione degli anticorpi contro il virus HIV e due anni dopo venne registrato negli Stati Uniti il primo farmaco attivo contro l'HIV, la Zidovudina (AZT).

1.1.3 CENNI DI EPIDEMIOLOGIA

I dati epidemiologici dei primi casi di AIDS dal 1981 a oggi sono stati fondamentali nel comprendere l'evoluzione della pandemia dell'infezione da HIV, specie nel quantificare le sue dimensioni e nel descrivere le modalità di trasmissione.

Secondo la GLOBAL HIV/AIDS RESPONSE, diffuso da UNAIDS, WHO e UNICEF, alla fine del 2011 circa 34 milioni di persone nel mondo erano affette da una infezione da HIV, di cui più di 3,4 milioni bambini. Guardando all'epidemiologia mondiale di HIV, ci sono significative differenze di tipo geografico: l'Africa sub sahariana detiene il primato di regione più colpita al mondo con quasi il 70% dell'intera popolazione infettata dal virus ^[6]. Nell'Europa orientale e nell'Asia centrale il numero della popolazione HIV positiva è aumentata da 650 mila nel 2001 a 1,5 milioni nel 2007. L'andamento dell'HIV in Europa centrale e occidentale tende a stabilizzarsi, con un leggero aumento dallo 0,2% nel 2001 allo 0,3% nel 2007. ^[7] Secondo il Notiziario dell'Istituto Superiore di Sanità nel 2011 in Italia sono stati diagnosticati 5,8 nuovi casi di HIV ogni 100.000 residenti.

L'introduzione della terapia HAART ha consentito, in Italia e negli altri paesi industrializzati, una riduzione sostanziale della morbosità e della mortalità associate ad HIV. Studi recenti hanno stimato che nei paesi industrializzati coloro che iniziano la terapia HAART hanno una speranza di vita dell'ordine di 40 anni. [8]

Malgrado ci siano stati progressi verso una terapia sempre più efficace, semplice da assumere e con bassa tossicità, i paesi in via di sviluppo hanno ancora un insufficiente accesso al trattamento HAART, pur presentando la più alta prevalenza della malattia da HIV (figura 1.1).

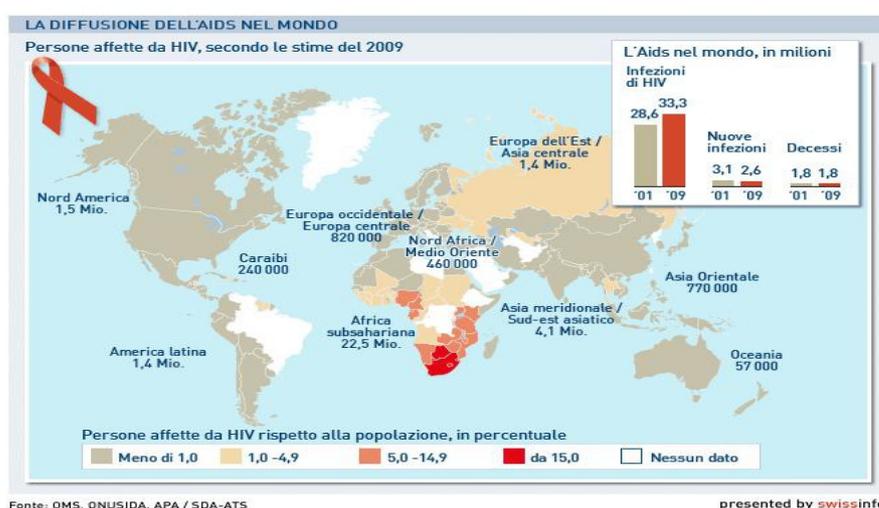


Fig.1.1: Diffusione dell'AIDS nel mondo stimata al 2009 dall'Organizzazione mondiale della sanità (OMS, o *World Health Organization*).

1.2 HIV

1.2.1 SOTTOTIPI DI HIV

I virus HIV di tipo 1 e tipo 2 appartengono al genere *Lentivirinae*, che con *Oncovirinae* e *Spumavirinae* costituiscono la famiglia *Retroviridae*. HIV-1 è il più virulento ed

infettivo, ed è l'agente eziologico della maggior parte dei casi di AIDS. HIV-2, meno virulento, induce una forma di malattia simile all'AIDS, ma con sintomi più lievi e un tasso di mortalità più basso. L'HIV-1 è prevalentemente localizzato in Europa, America e Africa centrale. HIV-2, invece, si trova per lo più in Africa occidentale e Asia. HIV-1 e HIV-2 si differenziano oltre che per la virulenza anche per l'organizzazione genomica, solo il 40% delle sequenze di HIV-1 e HIV-2 sono identiche. L'analisi filogenetica basata sulle sequenze del gene env ^[9] ha permesso di identificare tre gruppi in HIV-1: M, O e N. Il gruppo M, che da solo sostiene la maggioranza delle infezioni, è costituito da nove sottotipi (A-I). L'HIV-2 comprende sei sottotipi (A-F).

1.2.2 MORFOLOGIA DEL VIRIONE

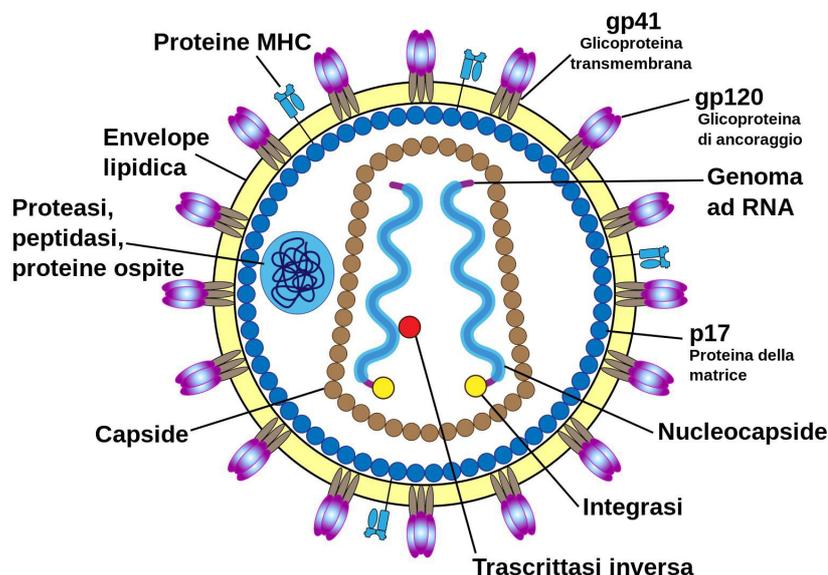


Fig.1.2: Struttura del virione di HIV

Il virione di HIV ha una forma quasi sferica ed un diametro di circa 100-120 nm. Esternamente è circondato da un involucro pericapsidico, detto envelope, che è costituito da un doppio strato fosfolipidico al quale sono ancorate le glicoproteine virali gp120 e gp41, implicate nell'interazione virus-cellula ospite. Il doppio strato fosfolipidico non è di natura virale, ma deriva dalla membrana citoplasmatica della cellula ospite che il virione acquisisce durante la gemmazione, nelle fasi finali della replicazione. ^[10] All'involucro pericapsidico segue uno strato di materiale elettrodenso, la matrice, costituita dall'assemblaggio di diverse unità della proteina 17 (p17). Il materiale genetico del virus è racchiuso nel core o nucleocapside del virione, che presenta una struttura cilindrica ed è costituita dalla proteina 24 (p24). All'interno del core si trovano i componenti necessari alla replicazione del virus quali il genoma, molecole di tRNA necessarie per l'inizio della replicazione, proteine non strutturali e i prodotti del gene pol, cioè gli enzimi trascrittasi inversa, integrasi, proteasi e ribonucleasi H (figura 1.2).

1.2.3 IL GENOMA DI HIV-1

L'HIV-1, come tutti i retrovirus, presenta un genoma formato da due molecole identiche di RNA monocatenario (RNA diploide a polarità positiva) di circa 9200 basi. Le due estremità 3' libere sono poliadenilate (poli A), mentre le estremità 5', parzialmente sovrapposte mediante legami a idrogeno, presentano un "capping" (Cap), costituito da residui di guanosina fosforilata e metilata. Ciascuna molecola di RNA genomico presenta, agli estremi, due sequenze di basi, delle quali la più esterna è ripetuta identica ai due estremi (sequenza R), mentre quella più interna è caratteristica rispettivamente dell'estremo 5' (sequenza U5) e dell'estremo 3' (sequenza U3). Quando il genoma di HIV viene retrotrascritto, ad opera dell'enzima trascrittasi inversa, il cDNA nascente subisce delle modifiche che lo rendono più lungo del genoma di partenza a causa della duplicazione ad entrambe le estremità delle regioni U3 ed U5. Le sequenze U3-R-U5,

presenti ad entrambi gli estremi della molecola, sono sequenze identiche non codificanti e altamente conservate denominate LTR (long terminal repeat), che contengono il promoter e l'enhancer per la trascrizione del provirus. ^[11] All'U3 di LTR 5' si trova infatti il sito promotore "TATA box" e la regione enhancer "CAT box", mentre all'R e U5 di LTR 3' vi è un segnale di poliadenilazione dell'mRNA, che determina la fine della trascrizione. Queste regioni, sono coinvolte nell'incapsidazione dell'RNA virale, nell'integrazione del virus nel genoma dell'ospite e nella regolazione del genoma virale stesso, poiché sono fondamentali per l'innescamento e l'amplificazione della trascrizione.

Procedendo dall'estremo 5' al 3', a valle rispetto ad U5, è possibile individuare il sito di legame del primer, *PBS* o *Primer binding site*, che funge da innescamento per la trascrizione inversa. A valle del PBS è presente una lunga sequenza avente due funzioni: fornire il segnale di incapsidamento Ψ , fondamentale per l'incorporazione dell'RNA nel virione e fornire il maggior sito donatore di splicing, utilizzato nella generazione di tutti gli mRNA subgenomici. Segue poi la regione codificante per i geni Gag-Pol, il sito di legame di Rev mentre il segnale di poliadenilazione costituisce l'estremo 3'.

La molecola di RNA contiene poi sotto-domini strutturali essenziali per il processo di trascrizione del genoma e per la circolarizzazione e integrazione dello stesso nel genoma della cellula ospite.

Il genoma di HIV-1 possiede 9 geni fondamentali, di questi tre, *gag*, *pol* ed *env*, sono geni strutturali tipici di tutti i Retrovirus, essenziali per codificare gli elementi necessari per la replicazione virale. Gli altri sei sono geni regolatori (*rev*, *tat*) e geni accessori (*nef*, *vpr*, *vpu* e *vif*), che permettono ad HIV una più complessa interazione con la cellula ospite. ^[12; 13]

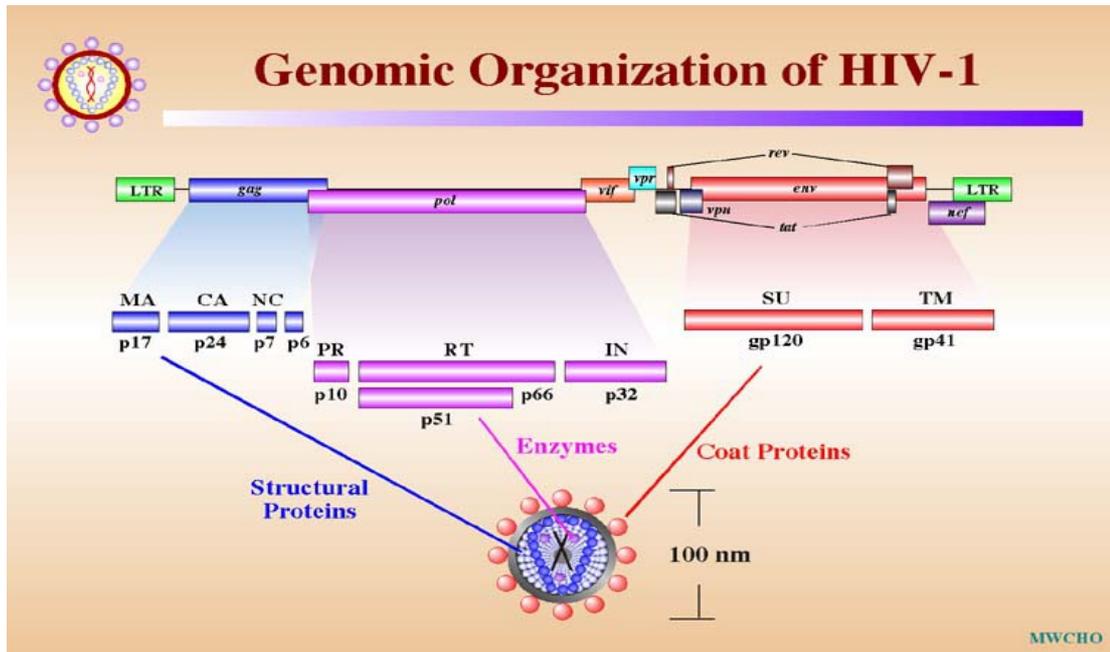


Fig. 1.3: Struttura del genoma di HIV-1 con i relativi prodotti di traduzione.

Caratteristica peculiare di HIV-1 è di possedere sequenze geniche con più schemi di lettura aperti (*ORF, open reading frame*). Ciò permette al virus di codificare una più ampia gamma di possibili proteine partendo da un frammento genico limitato.

1.2.4 PROTEINE STRUTTURALI DI HIV-1

I tre geni strutturali di HIV sono organizzati, dall'estremo 5' al 3', nel seguente ordine: *gag* (*group specific antigen*), *pol* (*polymerase*), *env* (*envelope*).

I geni *gag* e *pol* sono inizialmente trascritti in un unico mRNA, che viene tradotto in una

poliproteina di 180 kDa (p180), successivamente scissa da una proteasi virale in una proteina di 55 kDa (**p55**) e nelle tre proteine a funzione enzimatica, rappresentate da una **proteasi** virus-specifica (p10 o PR), da una **trascrittasi inversa** (p51/66 o RT) e da un'**integrasi** (p32 o IN).

Il prodotto del gene *gag*, la proteina **p55**, è tagliata dalla proteasi virale in quattro prodotti proteici: la p24, principale proteina capsidica, la cui funzione è di impacchettare il genoma virale nei virioni; la p17 o proteina di matrice, che si localizza tra il nucleocapside e l'envelope virionico, la p9 proteina del nucleocapside che si lega strettamente al genoma e infine la p7.

Per quanto riguarda i prodotti del gene *pol*, cioè le proteine virali a funzione enzimatica, la **proteasi** (PR o p10) è una proteina di 99 aa, che agisce in forma di dimero. Essa interviene nella processazione delle proteine virali a partire dai precursori poliproteici e gioca perciò un ruolo fondamentale nella maturazione della particella virale.

La proteasi agisce sia a livello della superficie cellulare durante il rilascio dei virioni, sia all'interno delle particelle già fuoriuscite dalla cellula, dove scinde i precursori gag e gag-pol,^[14] consentendo così la maturazione e l'infettività della particella virale.

La **trascrittasi inversa** (RT) è un enzima chiave per il ciclo replicativo di HIV, poiché permette la retrotrascrizione del genoma virale a RNA nella forma provirale a DNA.

RT è un eterodimero costituito da due subunità, la prima di 560 aa (p66) e la seconda di 440 aa (p51), entrambe derivate dalla poliproteina *pol*.

Le attività catalitiche (DNA-Polimerasica e Ribonucleasica) di RT sono svolte dalla subunità enzimatica p66, mentre la p51 è essenziale per stabilizzare il legame del primer del tRNA all'inizio della retrotrascrizione, per lo spostamento della catena e la fase di processazione durante la sintesi del DNA.

Il gene *env* è tradotto inizialmente in una proteina (p88) che viene rapidamente glicosilata, così da raggiungere il peso di 160 kDa (gp160), dal cui taglio proteolitico in fase di gemmazione derivano le due glicoproteine virus-specifiche di membrana, **gp41** e **gp120**.

Gp41 o TM (*Transmembrane*) è una proteina di 345 aa, situata nella membrana virale è coinvolta sia nell'ancoraggio di gp120 all'envelope, sia nel meccanismo di fusione tra la

membrana virale e quella della cellula ospite in seguito al legame con il recettore.

L'estremo N-terminale di gp41 è responsabile dell'ancoraggio di gp120 alla membrana grazie al legame con la regione COOH-terminale della stessa, mentre l'attività fusogena è mediata dalla presenza di una sequenza altamente apolare, ricca di glicine, che risiede in prossimità dell'estremo NH₂ e viene esposta in seguito al legame di gp120 al recettore CD4. ^[15] La glicoproteina gp41 è inserita nell'involucro lipidico pericapsidico, mentre la glicoproteina gp120 è ancorata con la porzione COOH all'estremo NH₂ terminale di gp41 ed è esposta sulla superficie del virione.

Gp120 o SU (*Surface*) è una proteina di 515 aminoacidi, la cui funzione principale è quella di promuovere il legame del virus alla membrana cellulare della cellula ospite mediante il riconoscimento della molecola CD4, considerata il principale recettore di HIV.

1.2.5 PROTEINE REGOLATRICI ED ACCESSORIE DI HIV-1

I geni regolatori principali *tat* e *rev* controllano l'espressione dei geni virali a livello trascrizionale e post-trascrizionale e sono il risultato della traduzione di mRNA originati da doppi eventi di splicing.

Il gene *tat* (*Trans Activator of Transcription*) codifica una proteina di 14Kd, costituita da 86 aa. Una volta sintetizzata, Tat entra nel nucleo cellulare e agisce principalmente come un potente attivatore della trascrizione del genoma provirale aumentando l'efficienza dell'RNA polimerasi II cellulare. ^[16]

La proteina Tat si lega ad una specifica sequenza (*Tar* o *Tat-responsive*) presente sugli RNA messaggeri virali nascenti, promuovendo la fase di allungamento della trascrizione del provirus di HIV, così da poter produrre l'intero trascritto.

I due geni regolatori essenziali *tat* e *rev*, assieme al gene accessorio *nef*, sono espressi precocemente e derivano dalla traduzione di RNA messaggeri di 2 kb. Il virus produce questi messaggeri mediante un processo di *splicing* dell'RNA di 9 kb, a sua volta

ottenuto dalla trascrizione del genoma provirale. Tat esercita un classico effetto di feedback positivo sulla trascrizione di HIV: una volta sintetizzata nel citoplasma, essa è in grado di entrare nel nucleo grazie alla presenza di un segnale di localizzazione nucleare (NLS) all'interno della propria sequenza, e di potenziare la trascrizione di HIV che porterà alla sintesi di altre proteine Tat. Il legame di Tat alle sequenze *tar* catalizza l'assemblaggio di proteine cellulari implicate nel controllo della proliferazione cellulare, quali la ciclina T1 e la chinasi CDK9.

Il gene *rev* (*Regulator of Virion Expression*), costituito da due esoni sovrapposti a quelli di *tat*, codifica una proteina di 116 aa (p19) a localizzazione prevalentemente nucleare. La proteina Rev ha una funzione cruciale nel regolare la sequenza di trascrizione dei geni virali. L'attività di Rev infatti media la transizione dalla produzione di trascritti primari sottoposti a più eventi di splicing (codificanti le proteine regolatorie e accessorie Tat, Rev e Nef), a trascritti prodotti da un singolo splicing e codificanti proteine strutturali. In questo modo Rev controlla il passaggio da un'infezione quiescente ad una replicazione virale attiva. ^[17] La proteina determina questa evoluzione trascrizionale legando il sito RRE (*Rev Response Element*), localizzato nella regione *env*. La regolazione svolta da Rev non è basata su un aumento della trascrizione degli RNA per le proteine strutturali, ma sul controllo del trasporto nucleo-citoplasma dei messaggeri virali, sull'inibizione dello splicing prima dell'esporto nucleare e sulla stabilizzazione dei trascritti corrispondenti attraverso il legame a RRE. Rev può anche inibire direttamente lo splicing bloccando l'ingresso nel nucleo di ribonucleoproteine responsabili dell'assemblaggio di uno spliceosoma funzionale.

Il gene *nef* (*Negative regulatory Factor*) codifica una proteina di 27 kDa (p27) a localizzazione prevalentemente citoplasmatica ed è la prima proteina ad accumularsi nelle cellule successivamente all'infezione con HIV. La sua funzione è quello di legare le molecole di CD4 e HLA-I e trasportarle dalla superficie cellulare verso i lisosomi, favorendone la degradazione. ^[18]

Il gene *vpr* (*Viral Protein R*) codifica una proteina di 96 aminoacidi (p15), incorporata

nelle particelle virali attraverso l'interazione con p6. La proteina, che sembra intervenire nella riattivazione dell'infezione latente, permette il rapido trasporto del complesso ribonucleoproteico al nucleo della cellula infettata, subito dopo l'uncoating del virus, grazie al suo NLS. ^[19] Vpr è responsabile anche del blocco del ciclo cellulare in fase G2 e dell'induzione della condensazione cromosomica, nonché della morte per apoptosi delle cellule T, in caso di una sua espressione prolungata.

Il gene *vif* (*Viral Infectivity Factor*), la cui sequenza nucleotidica si trova tra *pol* ed *env*, codifica una proteina di 192 aminoacidi del peso molecolare di 23 kDa (p23), importante per la produzione di virioni maturi altamente infettanti in cellule permissive e semipermissive. La proteina si trova sia nel citoplasma delle cellule ospiti, sia all'interno dei virioni liberi. ^[20] Essa viene inoltre regolata tramite modifiche post-trascrizionali.

Questa proteina, indispensabile per la produzione di virus infettanti, partecipa all'assemblaggio e alla maturazione delle particelle, intervenendo nell'elaborazione finale dei prodotti del gene *env*. La proteina è perciò essenziale per un'efficiente trasmissione del virus in cellule sane, mentre non influenza la diffusione cellula-cellula, che risulta solo ridotta in assenza del gene *vif*. ^[21]

Il gene *vpu* (*Viral Protein U*), codifica una proteina di membrana di 81 aminoacidi, del peso di 16 kDa (p16), con un dominio citoplasmatico C-terminale e uno idrofobico N-terminale di 24 residui. La funzione della proteina Vpu è quella di facilitare il trasporto del prodotto del gene *env* verso la membrana cellulare, ^[22] promuovendo la degradazione di CD4 responsabili della ritenzione, a livello del reticolo endoplasmatico, delle proteine Env neosintetizzate. La proteina, inoltre, regola negativamente l'espressione delle molecole MHC di classe I, mascherando così il virus al sistema immunitario e impedendo il riconoscimento e l'uccisione delle cellule infette da parte dei linfociti T citotossici.

1.2.6 CICLO REPLICATIVO DI HIV-1

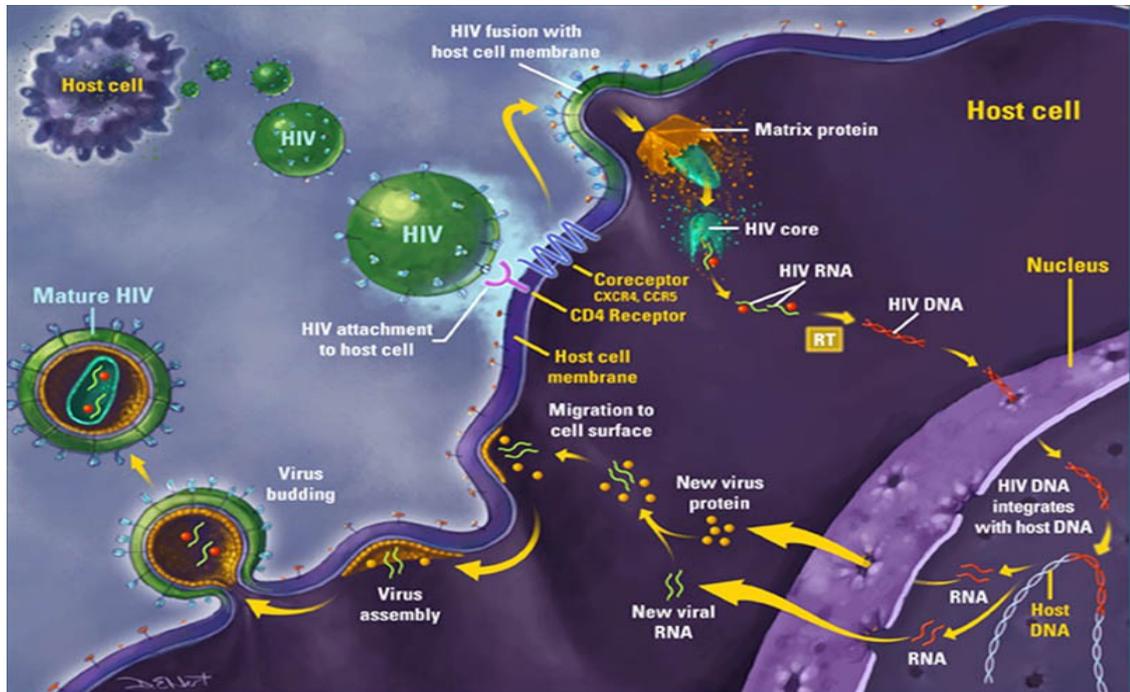


Fig.1.4: Ciclo riproduttivo di HIV

Il ciclo di replicazione dell'HIV-1 è suddiviso in diverse fasi: nella prima fase avviene il contatto del virus con la cellula bersaglio, la penetrazione del virus e la successiva scapsidazione ad opera di enzimi cellulari; nella seconda fase il genoma virale a RNA viene retroscritto in DNA il quale, insieme ad alcune proteine, viene traslocato nel nucleo cellulare dove avviene l'integrazione del DNA virale (*provirus*) nel genoma della cellula ospite; segue una terza fase di trascrizione del genoma virale in RNA messaggeri i quali, dopo opportuni eventi di *splicing*, vengono tradotti dall'apparato ribosomiale cellulare nelle proteine virali; nell'ultima fase le proteine virali vengono assemblate a formare nuovi virioni e la progenie gemma attraverso la membrana della cellula ospite.

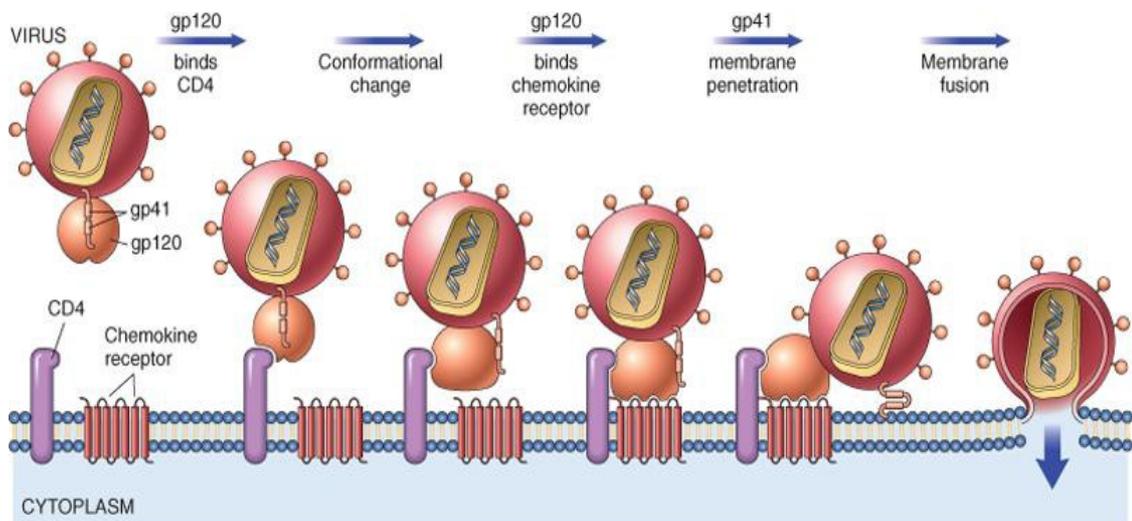
- **ADESIONE, PENETRAZIONE E UNCOATING**

Come per tutti i retrovirus, il ciclo replicativo di HIV comincia quando una particella virale si lega alla superficie esterna di una cellula ospite. La glicoproteina gp120 presente sull'envelope virale interagisce con la molecola CD4 (figura 1.5), una proteina transmembrana di 55kDa che rappresenta il recettore principale per il virus. Il CD4 si trova espresso principalmente su linfociti T helper, considerati il bersaglio d'elezione per HIV. Il virus inoltre è in grado di infettare anche altre tipologie cellulari quali monociti, cellule follicolari dendritiche, macrofagi alveolari, cellule della microglia, linfociti B, linfociti T CD8+, cellule endoteliali, cellule dell'epitelio intestinale e precursori emopoietici del midollo osseo.

Il legame di HIV con il CD4 non è tuttavia sufficiente per l'ingresso del virus nella cellula ospite, è necessario un secondo legame fra gp120 e un co-recettore presente sulla superficie della cellula bersaglio costituito principalmente da CCR5 nei macrofagi e CXCR4 nei linfociti T.

Il legame di gp120 con il co-recettore risulta in una modificazione conformazionale della glicoproteina virale che a sua volta causa l'esposizione e l'attivazione della proteina transmembrana gp41. ^[23]

La glicoproteina gp41 contiene 2 domini α -elica che in seguito all'attivazione di gp41 espongono la porzione N-terminale della proteina, detta peptide di fusione. Il peptide di fusione esposto è in grado di penetrare all'interno del doppio strato fosfolipidico della membrana citoplasmatica, consentendo così la fusione del pericapside virale con la cellula come conseguenza il nucleocapside virale viene rilasciato nel citoplasma della cellula ospite. ^[24;25]



© Elsevier. Kumar et al: Robbins Basic Pathology 8e - www.studentconsult.com

Fig.1.5: Ingresso di HIV-1 nella cellula bersaglio.

Una volta penetrato nella cellula, il core perde il proprio rivestimento proteico che viene degradato in un processo chiamato *uncoating* (svestimento); in questo modo si libera la parte centrale del virus che contiene il genoma ad RNA e gli enzimi virali.

- **TRASCRIZIONE INVERSA E INTEGRAZIONE**

La trascrizione inversa avviene nel citoplasma della cellula ad opera della Trascrittasi inversa. La molecola di DNA virale viene prodotta in tre fasi: nella prima fase avviene la sintesi di una catena di DNA complementare all'RNA virale (attività DNA polimerasica RNA dipendente della trascrittasi inversa), nella seconda fase avviene la degradazione della catena di RNA originaria (attività ribonucleasica della trascrittasi inversa) ed infine la terza fase che consiste nella sintesi della seconda catena di DNA complementare alla prima (attività DNA polimerasica DNA dipendente della trascrittasi inversa).

Al termine della retrotrascrizione il genoma virale è costituito da una duplice catena di DNA lineare contenente ad entrambe le estremità le sequenze LTR (Long Terminal

Repeats), associato alla proteina della matrice (p17), alla proteina Vpr e all'integrasi, forma il complesso di preintegrazione (PIC) che migra nel nucleo cellulare. In questa sede, grazie all'intervento dell'enzima virale integrasi, il DNA virale (provirus) viene inserito nel genoma cellulare.

Una volta integrato il DNA virale rimane associato al genoma della cellula ospite per tutta la durata della vita della cellula stessa, ed esso può rimanere trascrizionalmente inattivo, latente, e può persistere per mesi o addirittura anni, o andare incontro ad una attiva produzione di progenie virale.

- **TRASCRIZIONE E SINTESI PROTEICA VIRALE**

L'inizio della trascrizione del DNA provirale si verifica soltanto quando la cellula infetta viene attivata dall'esposizione ad antigeni, citochine e alcuni virus come CMV, HBV e EBV.

L'attivazione dell'apparato trascrizionale cellulare induce l'espressione di fattori di trascrizione che si legano alle sequenze LTR del provirus iniziando così la trascrizione del DNA virale. Vengono trascritti prima gli mRNA che codificano le proteine con funzione regolatrice, soprattutto Tat e Rev e solo in un secondo tempo si ha la trascrizione di mRNA che codificano le proteine strutturali necessarie alle successive fasi di assemblaggio e produzione di nuove particelle virali. Il passaggio dalla prima alla seconda fase dipende dall'azione della proteina Rev, che interviene nel trasporto degli mRNA dal nucleo al citoplasma (si veda pag 9). ^[26]

- **MATURAZIONE E GEMMAZIONE**

L'assemblaggio della particella virale di HIV si svolge a livello del citoplasma, dove l'aumentato livello di proteine e genomi virali, dà il via al processo. Le proteine che andranno a formare il capsido si riuniscono al di sotto della faccia interna della

membrana citoplasmatica, dove si trovano inserite la gp120, gp41 e p17. La proteina Vpu aumenta la velocità di montaggio dei componenti del virione e accelera il rilascio delle particelle virali dalla membrana della cellula ospite; la proteina Vif sembra invece essere in grado di aumentare l'infettività delle nuove particelle virali prodotte. La maturazione è lo stadio in cui i virus diventano infettanti. Tale fase implica variazioni strutturali nella particella virale, che possono derivare da tagli specifici delle proteine capsidiche per formare i prodotti maturi, da cambiamenti strutturali durante l'assemblaggio e da modificazioni interne, come la condensazione delle proteine virali col genoma del virus.

Durante il rilascio del virione dalla cellula le proteasi, tagliano il precursore derivato da *gag* nei prodotti finali. Lo stadio finale della maturazione delle particelle virali consiste nel rilascio delle particelle stesse dalla cellula ospite, processo che avviene per gemmazione dalla membrana cellulare, precedentemente modificata per l'inserimento delle proteine codificate dal gene *env*, attraverso la quale i virus acquisiscono l'envelope lipidico. ^[27]

1.3 INTEGRASI

1.3.1 STRUTTURA

La proteina virale Integrasi (IN), è costituita da 288 amminoacidi (32kDa) ed è ascrivibile alla superfamiglia delle polinucleotidil transferasi.

IN è costituita da tre domini strutturali: il dominio amino terminale (NTD), il core catalitico (CCD) ed il dominio carbossi terminale (CTD). [28; 29]

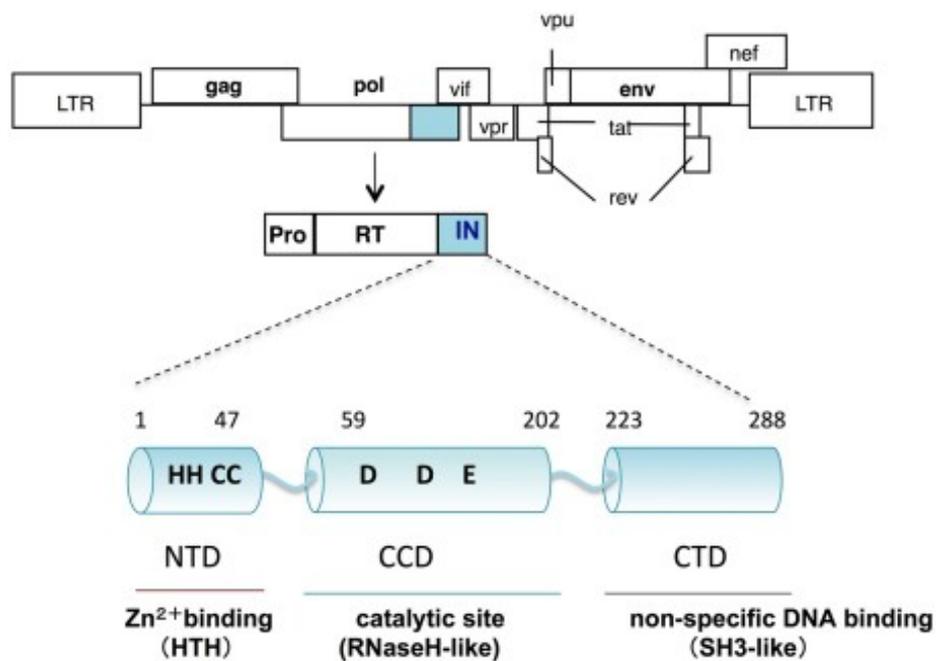


Fig.1.6: Domini dell'Integrasi e relative funzioni.

Il dominio **NTD** (aa 1-49) è strutturalmente organizzato in quattro α -eliche (*helix-turn-helix*), che costituiscono un tipico dominio zinc finger His2-Cys2 che coordina un atomo di Zinco. Diversi studi hanno dimostrato l'importanza di questo dominio nelle interazioni proteina-DNA e/o proteina-proteina, nella multimerizzazione dell'enzima, nella

stabilizzazione del folding di IN e nell'interazione con il co-fattore cellulare LEDGF/p75, necessario per il tethering e per la corretta integrazione del genoma virale nel genoma della cellula ospite. ^[30] Mutazioni a carico dei residui chiave del dominio zinc finger, la tetraide His2-Cys2, portano alla riduzione dell'attività enzimatica. ^[31]

Il dominio **CCD** (aa 50-212) costituito da una struttura α/β contenente 5 foglietti β fiancheggiata da 6 α -eliche, strutturalmente è simile a quello delle altre integrasi retrovirali. Nel sito catalitico è presente un loop, una regione flessibile che in presenza del DNA substrato assume una conformazione funzionalmente rilevante.

Il dominio catalitico rappresenta la porzione più conservata dell'integrasi e contiene una tripletta di amminoacidi altamente conservata in tutte le integrasi e le retrotransposasi. Tale tripletta, detta motivo DDE, è costituita da due residui di aspartato, rispettivamente in posizione 64 e 116 (D64 e D116) e da uno di glutammato in posizione 152 della struttura primaria di IN (E152).

Il funzionamento del sito catalitico dipende dalla presenza di due ioni divalenti Mg^{2+} in vivo, Mg^{2+} e Mn^{2+} in vitro, i quali sono coordinati dai residui della triade DDE. Il sito catalitico così costituito è essenziale per l'attività catalitica di IN tanto che l'assenza dei cofattori metallici o mutazioni a carico della triade aboliscono completamente il processo di integrazione. ^[32; 33]

All'interno del dominio CCD è presente, inoltre, un segnale di localizzazione nucleare che media l'importo al nucleo dell'IN nel contesto del complesso PIC. Inoltre alcuni residui presenti all'interno del dominio CCD sono coinvolti in vivo nell'interazione di IN con la proteina cellulare LEDGF (*Lens Epithelium-Derived Growth Factor*), interazione essenziale per l'integrazione di HIV. ^[34]

Il dominio **CTD** (aa 213-288) presenta una struttura simile al dominio SH3 ed è costituita da 5 foglietti β antiparalleli. ^[35] Dal punto di vista funzionale, il dominio CTD è coinvolto nei processi associativi quali la multimerizzazione, le interazioni con le proteine virali nel PIC e nella formazione di legami aspecifici con il DNA; ^[36] contiene infatti una regione a forma di doppia sella in grado di legare direttamente il DNA

attraverso una serie di residui contenuti nella sequenza del dominio. [37] Come per i domini NTD e CCD l'assenza del CTD abolisce l'attività catalitica dell'enzima.

Da un punto di vista della struttura tridimensionale delle integrasi, ad oggi, solo per il virus PFV (*Prototype Foamy Virus*) è disponibile una struttura completa. [38]

Questo è probabilmente dovuto a causa della scarsa solubilità delle IN. Sono state però determinate, tramite cristallografia a raggi X e spettroscopia NMR, le strutture cristalline che comprendono i singoli domini, e le strutture del dominio catalitico assieme al dominio N-terminale oppure assieme al dominio C-terminale. [39; 40]

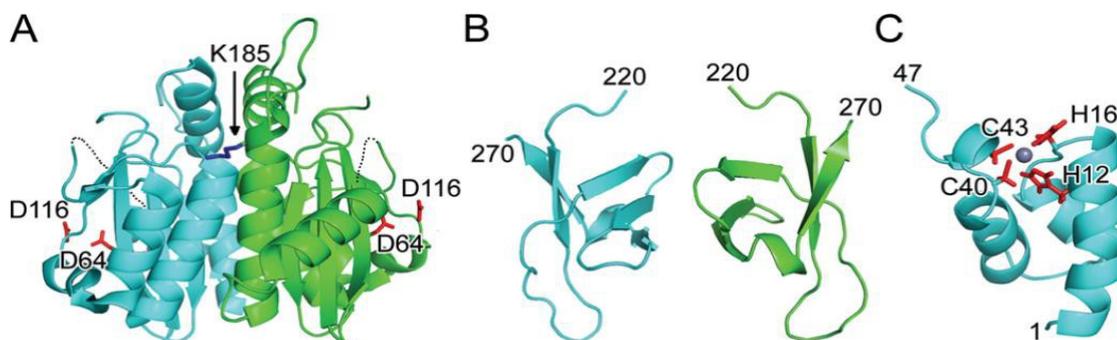


Fig. 1.7: Struttura dei domini dell'Integrasi di HIV-1 ottenuta mediante spettroscopia NMR. In A è visibile un dimero di CCD. Il pannello B mostra un dimero di CTD. Nel pannello C è visibile il dominio NTD di HIV-1 con lo ione zinco coordinato dal motivo HHCC.

Diverse evidenze sperimentali sembrano indicare che la struttura funzionale dell'IN è probabilmente multimerica, [41; 42] ed è stato ampiamente dimostrato come per una completa reazione di integrazione in vitro sia indispensabile la formazione di un tetramero in cui due dei quattro siti attivi sono coinvolti nella reazione di integrazione mentre gli altri due siti attivi fungono da supporto strutturale per il tetramero. [43]

1.3.2 FUNZIONE DELL'INTEGRASI

L'enzima Integrasi (IN) di HIV utilizza le sequenze LTR presenti alle estremità del DNA virale per catalizzare due reazioni chimiche distinte spazialmente e temporalmente, il *3' Processing* (3'P) e lo *Strand Transfer* (ST) entrambe necessarie per l'integrazione del DNA provirale nel cromosoma della cellula ospite. ^[44]

In seguito alla formazione nel citoplasma del complesso di preintegrazione (PIC), alla formazione del quale cooperano proteine virali e cellulari, ha inizio la fase di *3' Processing* durante la quale IN, in forma multimerica, rimuove i due nucleotidi 3' terminali (GT) di entrambi i filamenti del DNA virale, adiacenti al motivo CA altamente conservato, questa reazione produce quindi una molecola di DNA con estremità 5' sporgenti e 3'-OH. In seguito a *3' Processing* il PIC attivato viene traslocato nel nucleo con un meccanismo attivo ATP dipendente, attraverso i pori nucleari, dove avviene la reazione di *Strand Transfer*.

IN presente nel PIC lega il DNA cellulare, e promuove l'integrazione del DNA virale all'interno di una sequenza di poche paia di basi del DNA ospite, la lunghezza di tale sequenza è nota e specie specifica, 5 paia di basi nel caso di HIV-1; ^[45] di contro non è ancora chiaro se esista o meno una esatta sequenza nucleotidica sul DNA ospite, che funga specificatamente da sito di integrazione.

Dal punto di vista chimico nella reazione di *Strand Transfer*, IN catalizza l'apertura a singolo filamento del DNA della cellula ospite su entrambi i filamenti. IN inoltre media l'attacco nucleofilo delle estremità 3'-OH del DNA virale attivato dalle aperture a singolo filamento presenti sul DNA cellulare e media la trans-esterificazione del gruppo 5'-P del DNA cellulare. Il processo di integrazione è completato dalla rimozione di un dinucleotide al 5' del DNA virale e dal legame all'estremità 3' del DNA cellulare (*5' Processing*). ^[45] Il prodotto ultimo dell'integrazione è una molecola di DNA virale, il *provirus*, fiancheggiato da sequenze ripetute, gli LTR, con una duplicazione di 5 bp della sequenza genomica che fianeggia il sito di integrazione (figura 1.8).

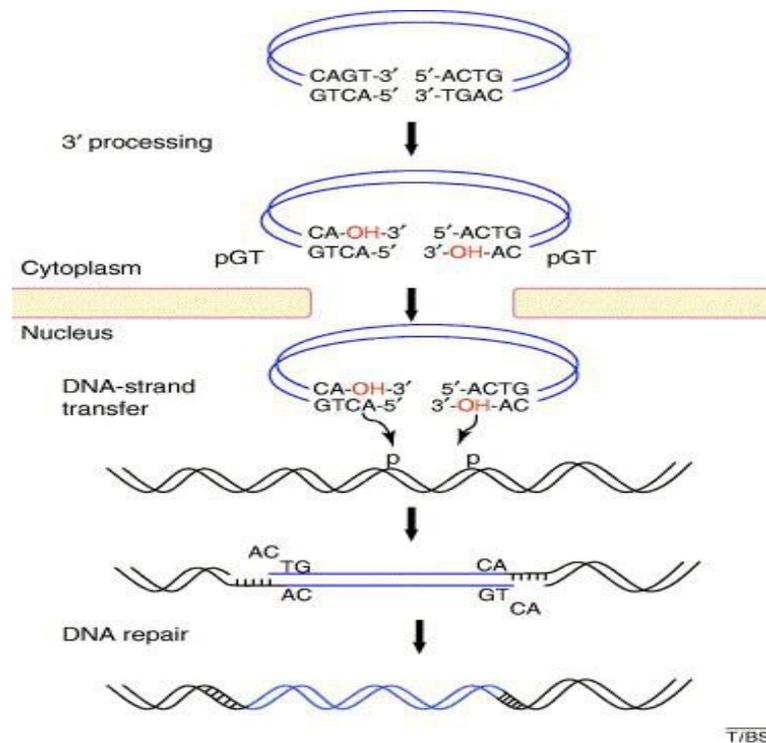


Fig. 1.8: Schema delle principali reazioni catalizzate dall'Integrasi di HIV-1: 3'Processing e lo *Strand Transfer*.

Le reazioni di 3' Processing e Strand transfer sono le principali reazioni catalizzate da IN nel corso dell'infezione; inoltre è stato dimostrato che in vitro l'IN è in grado di catalizzare anche altre reazioni:

- **Disintegrazione:** In presenza di un substrato di DNA che mima un prodotto incompleto di integrazione, l'integrasi può catalizzare il processo inverso dello Strand Transfer, noto come "disintegrazione": ^[46] IN è in grado di riconoscere il prodotto incompleto di integrazione e risolverlo in modo specifico clivando la porzione terminale del DNA virale e ricostituendo il sito di integrazione sul DNA bersaglio. Questo processo può essere catalizzato efficientemente sia da versioni full-length di IN che da

proteine tronche come l'IN¹⁻²¹² e l'IN⁵⁰⁻²⁸⁸ rispettivamente IN deleva del dominio CTD e IN priva del dominio NTD. Di contro le altre reazioni catalizzate in vitro da IN richiedono la forma completa di tutti e tre i domini.

- **Palindrome cleavage:** All'interno delle cellule infette si possono osservare tre forme di DNA virale: a) DNA lineare delimitato da due LTR, b) DNA circolare con una sola copia delle sequenze LTR e c) DNA circolare con due LTR giustapposti. Mentre il DNA lineare è il substrato immediato dell'enzima IN, è stato ipotizzato che le forme circolari del genoma virale, dette LTR-circles possano costituire un reservoir e venire linearizzate ed utilizzate in seguito da IN. È stato infatti dimostrato che in vitro IN è in grado di linearizzare i circoli 2LTR specificamente a livello della giunzione delle sequenze LTR. È stato inoltre dimostrato che in questa reazione il dominio CCD dell'IN, si lega in maniera specifica alle porzioni terminali di entrambi gli LTR, riconoscendo di fatto le sequenze bersaglio "naturali". Una volta riconosciute le sequenze virali bersaglio IN catalizza l'idrolisi a doppio filamento delle giunzioni LTR-LTR. Questa reazione avviene simmetricamente su entrambi i filamenti, con una forte preferenza di taglio dopo il dinucleotide CA, (la sequenza usata per la reazione di 3'P) e una minore preferenza dopo il dinucleotide TA, in direzione 5'-3'. La reazione di taglio avviene a livello di siti palindromici delle giunzioni LTR e richiede una forma full length di IN e la sua organizzazione in tetrameri. ^[47]

1.3.3 SITI DI INTEGRAZIONE

I siti riconosciuti dall'IN sul genoma virale sono chiamati siti *att* (*attachment sites*), sono regioni che agiscono in cis e risultano essenziali per l'integrazione virale. Da studi di mutagenesi è emerso che per il riconoscimento specifico sono sufficienti 11 o 12 bp rispettivamente alle estremità U5 e U3 degli LTR del DNA virale. Tale riconoscimento da parte dell'IN avviene in maniera indipendente e coordinata per ogni sito *att* anche se

l'enzima agisce probabilmente formando dimeri o tetrameri. ^[48] Numerosi studi hanno evidenziato che i nucleotidi conservati CA coinvolti nel 3'P giocano un ruolo fondamentale nel riconoscimento delle estremità virali da parte dell'IN, infatti mutazioni a livello di questa sequenza alterano l'attività dell'IN in vitro. Ad oggi, la determinazione di una sequenza minima necessaria per il riconoscimento specifico dell'IN rimane ancora oggetto di studio.

Per quanto riguarda l'IN, il dominio catalitico CCD si lega specificamente ai nucleotidi analoghi presenti a livello degli LTR virali mentre il dominio CTD dell'IN sembra essere coinvolto nello stabilizzare l'interazione con il DNA virale. ^[49] Sono stati elaborati diversi modelli virtuali per poter spiegare l'interazione del DNA virale con un multimerico di IN in quanto ad oggi manca il cristallo dell'intero enzima.

Da questi modelli è emerso che durante il posizionamento dell'IN sul DNA avviene un cambiamento conformazionale dei domini C-terminali che ruotano verso i domini catalitici. Ciò crea uno stretto canale che circonda il DNA virale tra il dominio C-terminale di una subunità e quello catalitico di un'altra subunità, perciò ciascuna LTR si lega ad un dimero. In questo modo le due estremità da processare si ritrovano l'una di fronte all'altra e ciò facilita l'attacco nucleofilo dell'estremità 3'-OH del filamento CA sul rispettivo filamento del DNA bersaglio. Esiste un ulteriore canale sulla superficie dell'enzima perpendicolare ai canali che accolgono le LTR e nel quale si inserisce il DNA target che in questo modo si avvicina ai residui del sito catalitico e alle estremità 3'OH delle LTR virali. ^[50]

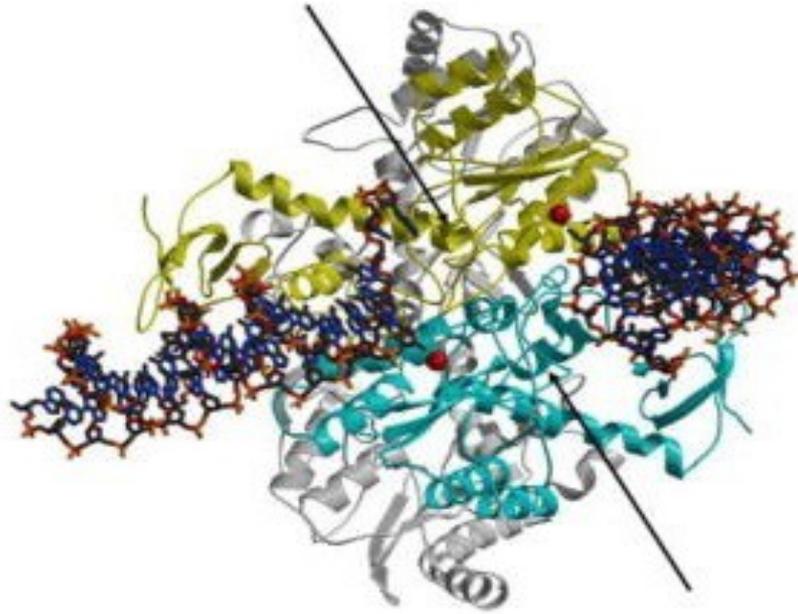


Fig. 1.9: Modello tetramerico di IN associato alle due LTR di HIV-1. Due subunità del tetramero sono mostrate in giallo e azzurro, le altre due in grigio. Il tetramero è ruotato per rendere evidente una delle due LTR.

L'analisi dei siti di integrazione *in vivo* indicano che HIV tende ad integrarsi in regioni trascrizionalmente attive nel DNA della cellula ospite e ciò può favorire l'espressione genica virale, facilitando così la produzione di particelle virali infettive, mentre l'integrazione in regioni della cromatina trascrizionalmente inattiva può sfavorire l'espressione genica virale, agevolando quindi, eventualmente, la latenza virale. ^[51] È stato inoltre visto che l'integrazione predilige sequenze palindromiche suggerendo che la simmetria del DNA target potrebbe stabilizzare l'organizzazione dei tetrameri di IN. ^[52]

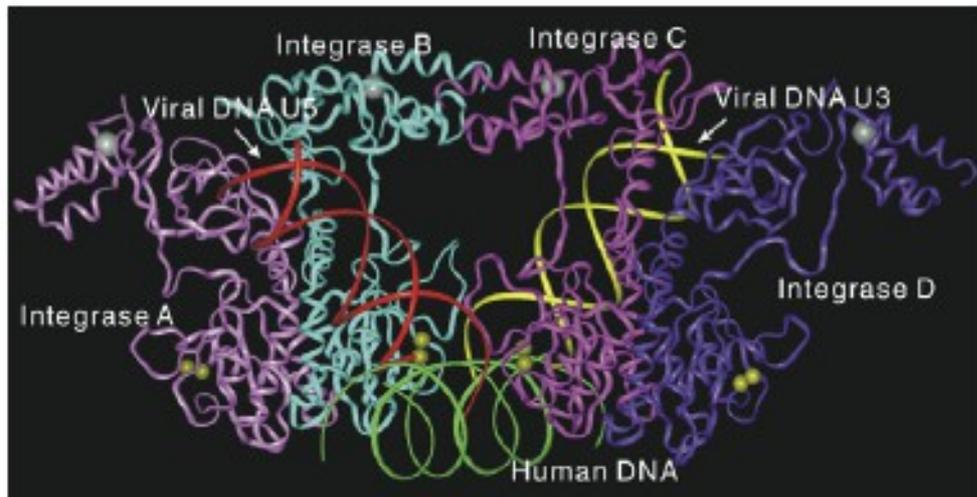


Fig.1.10: Tetramero di IN legante il DNA virale e umano.

1.3.4 COFATTORI

L'attività dell'IN in vivo è regolata da numerose proteine virali e cellulari. Tra questi abbiamo:

LEDGF/p75 (*Lens epithelium-derived growth factor*), proteina nucleare implicata nella regolazione trascrizionale, presenta nella regione C-terminale (residui 340-417) un dominio di legame per l'IN mentre nella regione N-terminale presenta un sito di legame che associa il PIC alla cromatina. In mancanza di LEDGF l'IN presenta una distribuzione citoplasmatica piuttosto che nucleare indicando un ruolo importante di questa proteina cellulare nella traslocazione di IN al nucleo e quindi, plausibilmente, anche dell'intero complesso PIC. ^[53]

p300 è un enzima istone acetil-transferasi (HAT) che si lega direttamente ad IN, sia in vivo che in vitro, acetilando tre residui di Lys nel CTD dell'IN. L'acetilazione promuove l'attività di *Strand Transfer* di IN e incrementa l'affinità di IN al DNA. Mutazioni puntiformi nei siti d'acetilazione, aboliscono la replicazione compromettendone la capacità d'integrazione. ^[54]

La proteina BAF (*Barrier-to-autointegration Factor*) è una componente del PIC che impedisce l'autointegrazione delle estremità del DNA virale all'interno dello stesso genoma virale. ^[55]

INI1 (*Integrase interactor 1*) è una componente del complesso proteico SWI/SNF della cellula. La sua funzione è rimodellare la cromatina mediante un meccanismo ATP dipendente e regolare l'espressione di alcuni geni eucarioti alterando l'interazione DNA-istone. In vitro è stato osservato che INI1 interagisce direttamente con IN di HIV e ne stimola l'attività. Particelle virali mancanti di tale fattore sono difettive per la retrotrascrizione indicando un potenziale ruolo di questa proteina nelle fasi precoci del ciclo vitale di HIV-1. ^[56]

1.4 TERAPIA FARMACOLOGICA E HAART

I farmaci antiretrovirali utilizzati per il trattamento del virus HIV non sono in grado di curare la malattia, ma possono impedire al virus di riprodursi. L'uso costante di questi farmaci permette di prolungare l'aspettativa di vita e migliorare la qualità della vita delle persone colpite dall'AIDS.

Sono sei le classi di farmaci che combattono l'HIV. Ciascuna classe ha un nome che deriva dal meccanismo d'azione contro il virus (figura 1.11).

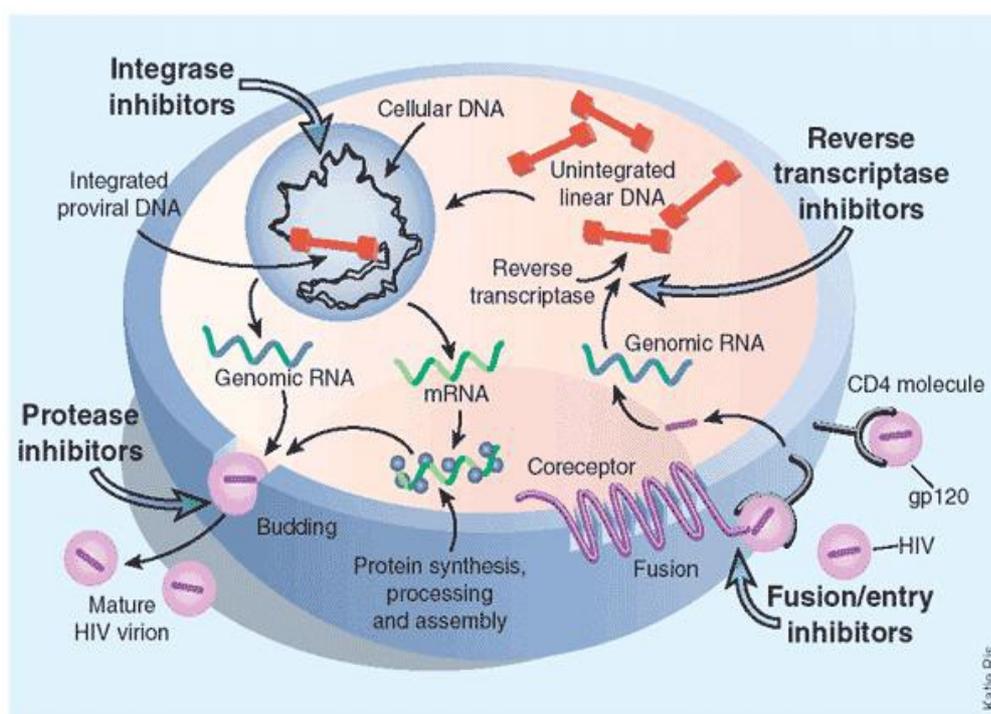


Fig.ra 1.11: In figura sono mostrati i bersagli degli attuali inibitori farmacologici di HIV-1

Inibitori nucleosidici della transcrittasi inversa (NRTI): La funzione di questa classe di farmaci è di bloccare la retrotrascrizione del genoma virale in copie di DNA. Gli NRTI vengono incorporati al posto della timidina dall'enzima RT e agiscono come

terminatori di sintesi della catena nascente di DNA virale.

Appartengono a questa classe i farmaci Zidovudina, Idovudina, Stavudina, Didanosina, Lamivudina, Emtricitabina , Abacavir e Tenofovir.

Inibitori non-nucleosidici della trascrittasi inversa (NNRTI): Gli NNRTI si legano all'enzima RT alterandone le capacità funzionali.

Appartengono a questa classe i composti Nevirapina, Efavirenz e Etravirina.

Inibitori della proteasi (IP): Gli IP bloccano la proteasi, l'enzima fondamentale per la maturazione delle nuove particelle virali, dando luogo alla formazione di particelle virali incomplete, prive di capacità infettante.

Appartengono a questa classe: Nelfinavir, Saquinavir, Tipranavir, Fosamprenavir Amprenavir, Atazanavir, Lopinavir, Darunavir e Ritonavir.

Inibitori della fusione (IF): Gli IF si legano alla glicoproteina gp41 e ne impediscono la funzione, bloccando l'ingresso dell'HIV nella cellula ospite.

A questa classe appartiene il composto Enfuvirtide.

Inibitori dell'Integrasi (II): Gli II bloccano il processo di integrazione del DNA virale nel genoma cellulare inibendo l'azione dell'enzima IN di HIV.

A questa classe appartiene il composto Raltegravir che agisce legandosi al complesso IN-DNA virale vicino all'estremità 3' bloccando lo strand transfer. ^[57]

Inibitori del corecettore CCR5: Gli antagonisti del corecettore CCR5 bloccano l'ingresso del virus dell'HIV nelle cellule sane dell'organismo.

Maraviroc è l'unico attuale rappresentante di questa classe di farmaci.

HIV presenta un elevato tasso di mutazione che gli consente di sviluppare selezioni clonali farmaco-resistenti rendendo inutile e dannosa la monoterapia, per cui la strategia vincente nella lotta all'AIDS è la terapia multifarmacologica o HAART (High Anti-

Retroviral Therapy) costituita dalla associazione di diversi farmaci (generalmente due NRTI con un inibitore NNRTI oppure con uno o due inibitori di proteasi.). [58; 59]

L'approccio HAART ha portato ad un significativo prolungamento della sopravvivenza e ad una marcata riduzione di infezioni opportunistiche e di decessi, tuttavia tale terapia presenta dei limiti come un costo molto elevato, uno schema posologico molto complesso, un'elevata tossicità. Inoltre non è in grado di eliminare i reservoirs virali latenti e quindi di eradicare definitivamente il virus dall'organismo.

HIV riesce a sopravvivere nell'ospite mediante la latenza pre e post integrazione. La latenza post-integrazione è causata da provirus inerti e metabolicamente inattivi che non sono accessibili né al sistema immunitario né ai farmaci; questo stato fornisce ad HIV un riparo sicuro nella cellula ospite e con opportuni stimoli è possibile riattivare uno stadio di latenza virale. L'esistenza di questi reservoir e le loro caratteristiche di estrema stabilità rendono irrealistiche le speranze di eradicare il virus con le terapie farmacologiche oggi in uso. E' quindi fortemente sentita l'esigenza di identificare nuovi approcci in combinazione con la terapia farmacologica in grado di eradicare il virus dai reservoirs in cui si nasconde.

1.5 APPROCCI ALTERNATIVI

Oggi la ricerca continua nello sviluppo:

- di nuovi farmaci meno tossici e più efficaci nel contrastare il virus, che inibiscano i complessi meccanismi di replicazione agiscano chiudendo le porte di entrata preferite nelle cellule del sistema immunitario;
- nella ricerca di nuovi approcci volti all'eradicazione dell'infezione in persone malate usando vari farmaci antivirali associati a sostanze immunostimolanti, inclusi alcuni prototipi di vaccini terapeutici;

- nel cercare ed utilizzare approcci alternativi ed innovativi, come la terapia genica, per prevenire e controllare la replicazione virale.

Il principio base della terapia genica consiste nell'inserzione, modificazione o rimozione di geni all'interno di specifici tipi cellulari e tessuti biologici, per trattare una malattia. Negli ultimi 15 anni differenti approcci di terapia genica sono stati testati sulle cellule ematopoietiche in vitro. Dai risultati di questi test è emerso che la terapia genica offre la possibilità non solo di prevenire e controllare la progressione dell'infezione da HIV, interferendo con la replicazione virale ma anche di prevenire la deplezione dei linfociti CD4+.

I primi approcci di terapia genica per il trattamento di malattie monogeniche si basavano sulla complementazione del gene endogeno mutato con l'introduzione casuale di un transgene funzionale. In questo modo però potevano verificarsi effetti secondari dati dall'attivazione di geni adiacenti al transgene. Ciò ha portato alla ricerca di una strategia per ottenere inserzioni in specifici loci evitando quelli casuali. ^[60]

1.5.1 GENOME EDITING

il *genome editing* è una tecnica di terapia genica che permette precise ed efficienti modificazioni genetiche mediante l'introduzione di una rottura del doppio filamento (DSB o *Duble-Strand Break*) all'interno di una sequenza genomica specifica, seguita dall'introduzione di modificazioni della sequenza stessa durante i meccanismi di riparo innescati dal taglio. Generalmente, i DSBs vengono indotti da nucleasi *zinc-finger*, capaci di riconoscere specifiche sequenze target.

Le cellule eucariotiche possono riparare i DSBs principalmente utilizzando due meccanismi: la ricombinazione omologa, *Homology Directed Repair* (HDR) e la giunzione delle estremità non omologhe, *Non-Homologous End Joining* (NHEJ). ^[61] Il meccanismo di ricombinazione omologa è attivo principalmente durante le fasi S e G2

del ciclo cellulare e utilizza il cromatidio fratello come stampo per il riparo del DNA. Il secondo meccanismo è il più diffuso nelle cellule eucariotiche ed è attivo in tutte le fasi del ciclo cellulare, principalmente durante le fasi G0, G1 e la fase S iniziale.

I prodotti ottenuti dalla ricombinazione omologa (HDR) sono:

- *Gene correction*: si ha quando il DNA donatore genera una variazione di una singola coppia di basi e ciò causa la modificazione dell'allele endogeno.
- *Targeted gene addition*: se il DNA donatore presenta una ORF o un transgene nella posizione corrispondente al sito di taglio, la sua sequenza viene trasferita al cromosoma mediante una via di *annealing* dipendente dalla sintesi del filamento.
- *Transgene stacking*: se il donatore presenta più transgeni successivi compresi tra le regioni di omologia essi vengono trasferiti nel cromosoma producendo una regione formata da geni disposti in successione. ^[62]

Il processo NHEJ invece, non richiede un DNA donatore e i suoi esiti possono portare a:

- *Gene disruption*: si ha con il processo di ligazione delle estremità, ciò può causare delle piccole inserzioni o delezioni con conseguente acquisto o perdita di informazione genetica nel sito di taglio.
- *Tag ligation*: se viene fornito un oligonucleotide a doppio filamento con estremità compatibili a quelle liberate dalla rottura esso verrà inserito nel cromosoma.
- *Large deletion*: la produzione simultanea di due DSBs sullo stesso cromosoma può portare alla delezione della sequenza di DNA compreso tra i due siti di rottura.

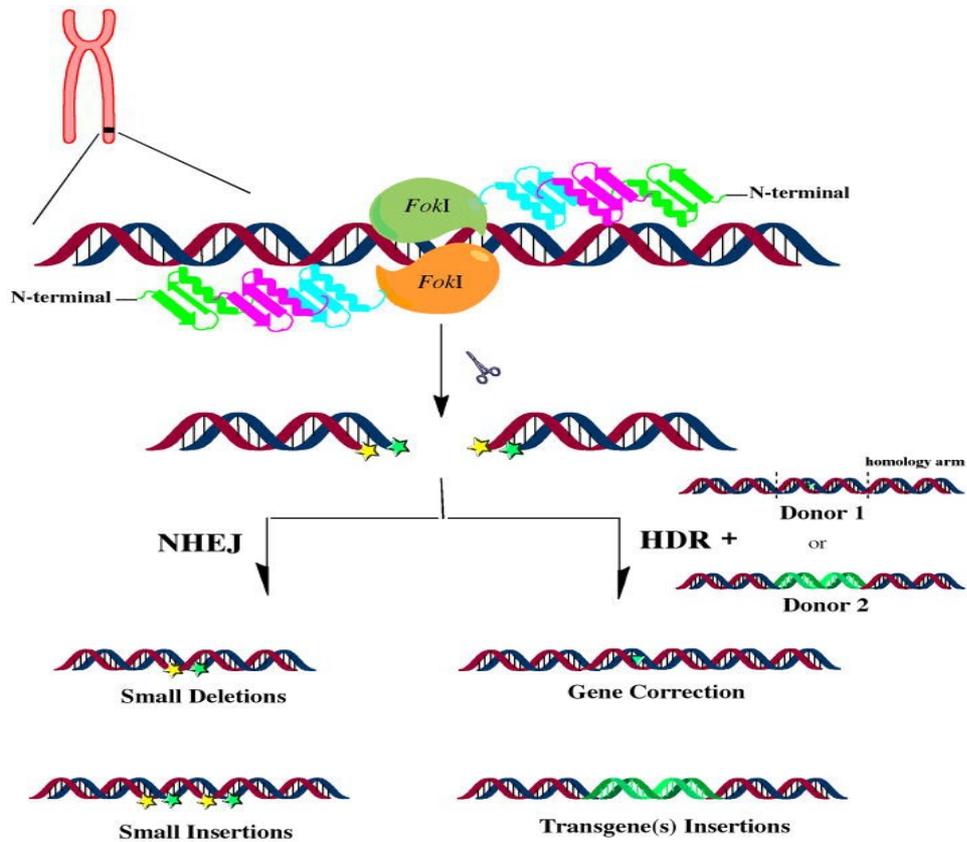


Fig. 1.12: Tipologie di *genome editing* rese possibili utilizzando le nucleasi zinc finger

1.5.2 NUCLEASI ZINC FINGER (ZFN)

L'approccio di terapia genica basato sulle nucleasi zinc finger, ad oggi, sembra essere tra i più promettenti per il potenziale trattamento di pazienti infetti da HIV.

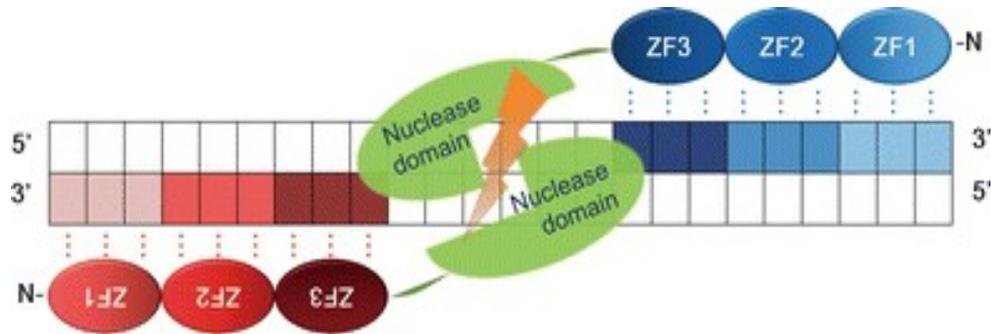


Fig.1.13: Rappresentazione di un dimero di ZFN legata al DNA target. Ogni ZFN presenta il dominio catalitico di FokI legato a tre zinc-finger .

Le nucleasi *zinc finger* sono composte da un dominio di taglio derivante da una subunità endonucleasica, spesso quella dell'enzima di restrizione *FokI*, fuso ad un dominio di legame al DNA, formato da moduli Cys2-His2 con struttura a “dito di zinco”.^[62] Tali moduli sono capaci di legare tre o quattro nucleotidi adiacenti (figura 1.13) e il legame avviene attraverso alcuni residui amminoacidici che contattano direttamente il solco maggiore del DNA. Sostituendo tali amminoacidi, è possibile modificare la specificità del singolo modulo. In natura *FokI* opera come un dimero: affinché il taglio possa avvenire i due domini di taglio devono disporsi in maniera antiparallela, ed il DNA da idrolizzare deve trovarsi tra essi (Figura 1.14). Per disegnare ZFN è quindi necessario disegnare accuratamente i domini di legame al DNA, in modo che la distanza che li separa consenta la formazione del dimero dei domini catalitici.



Fig.1.14: Dimero di ZFN con quattro domini zinc-finger.

Le prime ZFN contenevano tre moduli zinc-finger per un DNA target di 9 bp totali permettendo al dimero di ZFN di avere una specificità di 18 bp totali. Studi più recenti hanno confermato che è possibile ingegnerizzare le ZFN con 6 moduli zinc-finger in modo da ottenere una maggiore specificità per il DNA target ed aumentare il sito di riconoscimento da 18bp a 36bp, ovvero una sequenza abbastanza lunga da ricorrere una sola volta nel genoma umano.

Approcci di genome editing con ZFNs sono stati adottati in vitro, con ottimi risultati, per modificare il genoma umano in maniera sito-specifica, attivando il meccanismo di riparo della ricombinazione omologa; ^[63] queste proteine chimeriche sono state applicate anche in vivo in modello animale per l'emofilia B, ^[64] in cui la mutazione nel gene del Fattore IX è stata corretta direttamente nelle cellule epatiche del topo, in modo da ottenere i livelli circolanti di tale molecola capaci di fornire un beneficio terapeutico (>1-5%).

Considerando i grandi successi ottenuti con sperimentazione animale, le ZFNs sono state applicate per la terapia genica ex vivo in cellule staminali ematopoietiche per inattivare il gene umano CCR5. La distruzione della sequenza codificante questo gene genera mutanti non funzionali che rendono la cellula resistente ad HIV. ^[65]

Negli approcci appena descritti la ZFN, viene introdotta come proteina o come unità genetica trascrizionale, in cellule staminali ematopoietiche in maniera transitoria:

l'espressione stabile molto probabilmente potrebbe causare genotossicità; la nucleasi quindi deve esser rimossa dopo un singolo step di mutagenesi. [65]

1.5.3 FOKI

L'enzima FokI è una proteina monomerica di 66,4 kDa, isolata da *Flavobacterium Okeanokoites*, appartiene alla classe IIS delle endonucleasi di restrizione.

Studi proteolitici, hanno dimostrato che FokI è costituita da due domini strutturali distinti connessi da un linker flessibile: un dominio N-terminale di 41 kDa per il riconoscimento del DNA e un dominio C-terminale di 25 kDa nucleasico che taglia il DNA in maniera aspecifica in presenza di Mg^{2+} . [66]

FokI esiste in soluzione come monomero, ma per poter idrolizzare entrambi i filamenti di DNA è necessaria la dimerizzazione dei domini catalitici di due monomeri di FokI in modo da ottenere il corretto allineamento dei due domini catalitici e permettere la scissione dei legami fosfodiesterici. [67]

L'enzima FokI riconosce la sequenza asimmetrica 5'-GGATG ed effettua tagli sfalsati: 9 nucleotidi dalla sequenza di riconoscimento del filamento senso e 13 nucleotidi dalla sequenza del filamento antisenso (figura 1.15).



Fig. 1.15: Sequenza riconosciuta da FokI e relativi siti di taglio (freccie).

La caratteristica struttura di FokI viene sfruttata per creare proteine chimeriche in cui si va a sostituire solo il dominio DNA binding di FokI con altre sequenze di legame al

DNA con attività di riconoscimento sequenza-specifica.

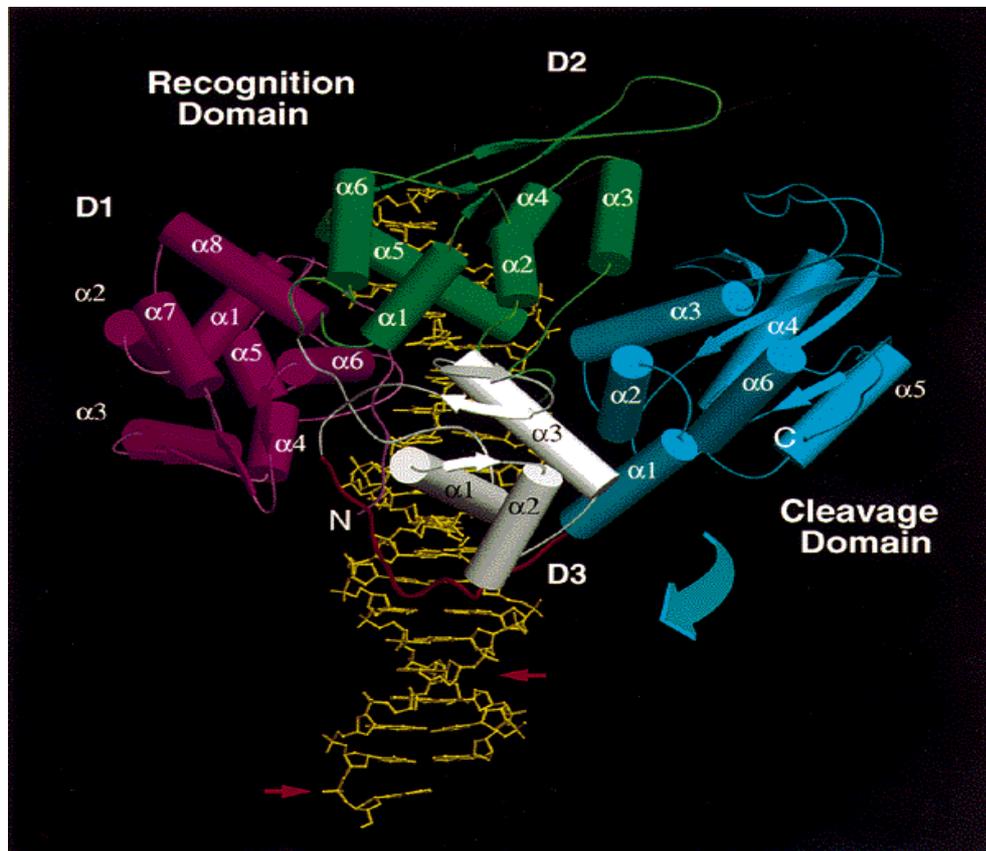


Fig1.16: FokI legato al DNA. Il dominio di riconoscimento consiste di 3 sottodomini D1 (in lilla), D2 (in verde) e D3 (in bianco). Il dominio di taglio (in blu), il linker (in rosso) e il frammento di 20 bp di DNA (in giallo). Le due frecce rosse indicano i siti di taglio. La freccia celeste mostra il movimento proposto per il dominio di taglio verso il suo sito bersaglio. ^[68]

PARTE SPERIMENTALE

2. OBIETTIVO DELLA RICERCA

Dalla scoperta del virus HIV ad oggi, sono stati fatti grandi passi in avanti nelle conoscenze sulla patogenesi della malattia, nelle procedure diagnostiche e negli interventi terapeutici.

L'avvento della terapia HAART, basata sul trattamento dei pazienti HIV positivi con farmaci antiretrovirali usati in combinazione, ha permesso di sopprimere la replicazione e prevenire la progressione in AIDS. Nonostante questo, l'AIDS resta ancora ai nostri giorni una malattia estremamente severa in quanto le terapie farmacologiche non eradicano il serbatoio di HIV dall'organismo infetto; oltretutto il trattamento prolungato e persistente con HAART determina la comparsa di ceppi HIV resistenti, rendendo quasi totalmente inefficaci i trattamenti farmacologici.

La grande variabilità degli antigeni di HIV rende al momento ancora molto complesso l'allestimento di un vaccino efficace, per cui la ricerca di terapie alternative come il *genome editing* può rappresentare un grande passo avanti per il trattamento della patologia.

E' in questo ambito che si colloca il progetto a cui ho lavorato principalmente in questi tre anni, cioè la creazione di una strategia per eradicare il provirus di HIV integrato nel genoma della cellula ospite.

A questo scopo è stato ideato un enzima chimerico tra la proteina IN di HIV e il dominio catalitico dell'enzima di restrizione *FokI*: l'enzima è stato disegnato per riconoscere specificamente le sequenze LTR presenti sul genoma del provirus integrato e conseguentemente per idrolizzare con un taglio a doppio filamento il DNA adiacente al

sito di riconoscimento.

Il dominio della chimera responsabile per il riconoscimento ed il legame delle sequenze LTR di HIV è l'enzima IN di HIV, che in cellule infette media l'integrazione del cDNA virale nel cromosoma della cellula ospite. Il dominio catalitico dell'enzima, deputato alla idrolisi a doppio filamento del DNA è costituito dalla subunità catalitica di *FokI*, che in vivo è in grado di idrolizzare il DNA a prescindere dalla sequenza legata dal dominio di riconoscimento del DNA. I due domini sono stati uniti da una breve sequenza di amminoacidi denominata Linker3 che conferisce flessibilità alla struttura.

L'obiettivo principale che ci siamo prefissati è stato quello di creare diverse varianti della chimera IN-L3-FokI, ottenute sia con mutazioni sito specifiche che con mutazioni random e di selezionare le varianti ottenute sottoponendole a sistemi di selezione in vivo.

Questo lavoro porterà a definire quindi se l'IN di HIV può essere riprogrammata a catalizzare una nuova funzione mediante la sostituzione dell'attività del proprio dominio catalitico con quello di FokI.

3. MATERIALI E METODI

3.1 COLTURE BATTERICHE

3.1.1 CEPPI BATTERICI

Per la propagazione di DNA plasmidico, per l'espressione di proteine ricombinanti e per la selezione in vivo sono stati utilizzati rispettivamente i seguenti ceppi batterici di E.coli:

TOP10 (Invitrogen), **BL21 starTM (DE3) One Shot** (Invitrogen) e **BW25.141** (Coli Genetic Stock Centre, Yale University).

Il genotipo del ceppo **TOP10** è *F-mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ-*. Il ceppo presenta una mutazione missenso del gene *recA*, utile per evitare la ricombinazione casuale, e una mutazione del gene *endA* che codifica una Dnasi. Le due mutazioni permettono un'ottima resa di trasformazione.

Le cellule batteriche **BL21 (DE3)**, caratterizzate dal genotipo *F- ompT hsdSB (rB-mB-) gal dcm rne131 (DE3) lacUV5 T7RNAP-tetA*, sono state utilizzate per l'espressione delle proteine ricombinanti. Questi batteri presentano il gene della RNA polimerasi T7 sotto il controllo del promotore lacUV5 e per l'espressione necessitano induzione mediante IPTG.

Inoltre contengono un gene mutato (*rne131*) che codifica un enzima Rnasi E tronco, incapace di degradare l'mRNA e ciò permette una maggiore stabilità degli RNA messaggeri. Infine non contengono le proteasi *lon* e *OmpT* e ciò riduce la degradazione di proteine eterologhe espresse dai batteri. A causa dell'alta attività della RNA polimerasi

T7 e una grande stabilità dei mRNA anche nelle cellule non indotte vi è un livello di espressione basale del gene di interesse.

E.Coli **BW25.141** realizzato dalla Coli Genetic Stock Center, Yale University, possiede una mutazione missenso, la G720A del gene *recA1*, che permette una maggiore stabilità genetica rispetto ai *wild-type* limitando la ricombinazione genica. Questo ceppo è in grado di degradare plasmidi lineari.

3.1.2 TERRENI DI COLTURA

Il terreno di coltura utilizzato per la crescita delle cellule batteriche di E. coli è il Luria Bertani (LB), che è composto da:

- 10g/L di Tryptone o bacto-tryptose
- 5g/L di Yeast extract
- 5g/L di NaCl
- 1g/L di glucosio
- 1ml/L di NaOH 1M

Per ottenere terreni solidi è stato aggiunto alla soluzione Agar all'1% (10g/L) mentre per la selezione in vivo, al terreno solido è stato aggiunto l'induttore del gene tossico *ccdB*: L-arabinosio 100mM (15g/L). Dopo aver portato a volume con acqua distillata, il terreno è stato sottoposto a sterilizzazione in autoclave.

3.1.3 PREPARAZIONE DEI BATTERI ELETTROCOMPETENTI

Per rendere i batteri elettrocompetenti è stato effettuato un pre-inoculo, O/N a 37°C, della coltura batterica d'interesse in LB in presenza dell'opportuno antibiotico nel caso di batteri trasformati in precedenza con plasmidi di interesse, o in assenza di antibiotico in caso contrario. Il giorno seguente parte della coltura è stata inoculata in un volume maggiore di LB (il rapporto inoculo : pre-inoculo è di 1:50) in presenza o meno di antibiotico (si veda sopra) ed è stata messa a crescere a 37°C fino al raggiungimento della densità ottica ottimale, (0,8-1.0 OD). I batteri così ottenuti sono stati centrifugati per 1h a 4000rpm e il pellet batterico è stato risospeso, in un volume pari a quello di partenza, in acqua sterile e fredda. Il pellet batterico è stato sottoposto a ripetuti lavaggi dimezzando ogni volta il volume di risospensione dell'acqua. Per l'ultimo lavaggio è stata usata una soluzione di glicerolo al 10% sterile e freddo. Infine il pellet, risospeso in glicerolo al 10%, è stato aliquotato in eppendorf, congelato istantaneamente in azoto liquido e conservato a -80°C.

3.1.4 TRASFORMAZIONE DEI BATTERI PER ELETTROPORAZIONE

I diversi costrutti plasmidici ottenuti sono stati introdotti in batteri competenti mediante elettroporazione ad alto voltaggio. Le condizioni di utilizzo dell'apparecchio (*GenePulser^R II Electroporation System, BioRad*) sono state:

Capacitanza: 25uF

Voltaggio: 25KVcm⁻¹

Resistenza: 200 Ohm

Dai 3-6 ng di DNA sono stati aggiunti a 50 µl di batteri nell'apposita cuvetta da 0,2 cm (spazio fra gli elettrodi) e quest'ultima è stata inserita nella camera dell'elettroporatore.

Dopo il passaggio della corrente elettrica è stato aggiunto 0,8ml di terreno LB e i batteri trasformati sono stati lasciati a 37°C per 45 minuti in agitazione per permettere loro di riattivare il metabolismo. Al termine dell'incubazione, volumi di 50-100µl di cellule trasformate sono state seminate su piastre contenenti terreno LB-agar e l'opportuno antibiotico per la selezione. In determinate condizioni, nella fase di incubazione, è stato aggiunto IPTG. Le piastre sono state poste capovolte a 37°C e incubate O/N.

Per quanto riguarda i batteri doppio trasformanti sono stati piastrati su terreno solido contenente gli opportuni antibiotici e l'arabinosio, necessario per indurre l'espressione dei geni del plasmide reporter.

3.1.5 PURIFICAZIONE DEL DNA PLASMIDICO DA *E. Coli*

I trasformanti ottenuti sono stati inoculati individualmente in LB contenente l'antibiotico selettivo e lasciati a crescere tutta la notte a 37°C. Si procede quindi all'estrazione del DNA plasmidico seguendo le indicazioni del kit commerciale “*Pure yield*TM *plasmid Miniprep system*” (Promega) o del kit “*NucleoBond*^R *Xtra Midi*” *Nucleic Acid and Protein Purification* (Macherey-Nagel).

L'integrità e la concentrazione del DNA purificato sono state verificate mediante elettroforesi su gel di agarosio.

3.2. BIOLOGIA MOLECOLARE

3.2.1 PLASMIDI

I plasmidi usati in questo progetto sono stati:

- pUC19 (New England Biolabs)

Vettore di 2686 bp che presenta il promotore lacZ, il gene per la resistenza all'ampicillina e l'origine di replicazione derivata da pMB1 con una mutazione singola puntiforme che lo rende un high copy number (500 copie/cell).

In questo plasmide è stata clonata la sequenza della chimera ibrido IN-L3-FokI.

- pCR8/GW/TOPO (Invitrogen)

Il vettore pCR8, 2817 bp, presenta il gene per la resistenza alla spectinomicina e l'origine di replicazione di pMB1 che ne conferisce caratteristica di plasmide high copy number. Tale vettore è stato utilizzato per la realizzazione dei plasmidi bersaglio contenenti gli LTR virali di HIV-1.

- pACYC184

È un vettore low copy number (10-20 copie /cell) di 4245bp. Contiene l'origine di replicazione di p15A che permette una maggiore stabilità in doppi trasformanti che esprimano due proteine ricombinanti simultaneamente, e la resistenza a cloramfenicolo e tetraciclina.

Questo plasmide è stato utilizzato per la realizzazione del vettore recante il gene lacY.

3.2.2 OLIGONUCLEOTIDI

Le sequenze nucleotidiche dei primer utilizzate in questo lavoro, riportate nella tabella 3.1, sono state sintetizzate dalla ditta IDT (Leuven, Belgio) e risospese in TE buffer 1X sterile pH8 ad una concentrazione di 5µM.

Alcuni oligonucleotidi includono nella loro sequenza siti di riconoscimento per enzimi di restrizione, facilitando in tal modo le successive reazioni di digestione e di ligazione del prodotto di PCR in un plasmide reso lineare con gli stessi enzimi.

NOME OLIGO	SEQUENZA 5'-3'	BP	NOTE
Fw Int D116A	GCCAGTAAAAACAATACATACAGCAATGGCAGCAATTCACC AG	45	D116A
Rev Int D116A	CTGGTGAAATTGCTGCCATTGCTGTATGTATTGTTTTACTGGC	45	D116A
Linker 3	<u>CCGCGGGGAGGAGGAGGAAGCGGTGGAGGAGGCAGCGGGC</u> <u>C GGTGGTAGCGCATGCTTTGGAGGTACC</u>	69	SacII, SphI, KpnI
Rev L3	<u>GGTACCTCCAAAGCATGCGTACC</u>	24	KpnI
Fw Fok	<u>GCCAGCATGCAAGCAACTAGTCAAAAGTGAAGTGGAGGAGA</u>	41	SphI
Rev FokI	<u>GGGGGGTACCTTAAAAGTTTATCTCGCCGTTATTA</u>	35	KpnI
Fw delCTD	GTAGACATAATAGCAACAGACATAAACTAAAGAACCGCGGG GAGGAGGAGGAAGC	57	
Rv IN- NoCTD	CTTTAGTTTGTATGTCTGTTGCTATTATGT	30	
Fw del2- Fok	CAAGAAGAAAAGCAAAGATCATTAGGGATCCGCGGGGAGGAG GAGGAAGCGGTGGAGGAG	60	
RvIN- del2	ATCCCTAATGATCTTTGCTTTTCTTCTTG	29	
att0	<u>GACCGGTTGTGGAAAATCTCTAGCAGTACTGGAAGGGCTAATTT</u> <u>ACTGTCGACGTC</u>	56	AgeI, Sall
att20	<u>GACCGGTTGTGGAAAATCTCTAGCAGTGTAGTATATGCTTGTGTT</u> <u>CCACTGGAAGGGCTAATTTACTGTCGACGTC</u>	76	AgeI, Sall

att40	<u>GACCGG</u> TTGTGGAAAATCTCTAGCAGTGTAGTATATGGTCATTC ACAAGACTCTAGTCTTGTGTTCCACTGGAAGGGCTAATTTACTG <u>TCGACGTC</u>	96	AgeI, Sall
RevLat3p UC	GGAATT <u>GGTACC</u> TTAAAAGTTTATCTCGCCGTTATTAAATTTCCG	45	KpnI
Fw Lat3 pUC	GGTTTAGTCGACATTTTTAGATGGAATAGATAAGGCACAAGA	42	Sall
IntDel2 Fw	CCTAGGGATTACTAGAAACGAAAAGAAGAAC	31	
IntDel2 Rev	CGCCCCTCCTCCTCCTTCGC	20	
IN FW	TTTTTAGATGGAATAGATAAGGCACAA	27	
IN REV Del1	GAGCTCATGCCGCGGACCTGCCATCTGTTTTCCAT	35	
IN REV Del2	GAGCTCATGCCGCGGATCCCTAATGATCTTTGCTTT	36	
IN REV Del3	GAGCTCATGCCGCGGTGGCACTACTTTTATGTCAC	35	

Tabella 3.1: Oligonucleotidi utilizzati per le clonazioni. I diversi siti di restrizione inseriti sono sottolineati. Codoni usati per produrre mutagenesi sono sottolineati e riportati in rosso.

3.2.3 AMPLIFICAZIONE MEDIANTE PCR

Il software Oligo (Molecular Biology insight, Inc. Cascade, CO, USA) è stato utilizzato per la progettazione e l'ottimizzazione delle sequenze di tutti gli oligonucleotidi utilizzati in questo studio; inoltre ha fornito una guida preziosa nella scelta delle condizioni ottimali di amplificazione. Per l'amplificazione è stato utilizzato il kit PerfectTAQ plus (5Prime) seguendo le indicazioni fornite dal produttore. Le condizioni generali sono indicate in tabella 3.2, seguite da programma di PCR.

	Concentrazione iniziale	Concentrazione finale
Buffer	10x	1x
dNTPs	2mM	0,2 mM
Primer mix	5 μ M	0,5 μ M
H2O		A volume
Taq	5U/ μ L	0,025U/ μ L
MgCl ₂	25mM	1,5mM
DNA template		1 ng

PROGRAMMA:

Ciclo 1:	95°C	x 3 min	
Ciclo 2:	95°C	x 30 sec	} 30 cicli
Ciclo 3:	56-60°C	x 30-40 sec	
Ciclo 4:	72°C	x 1Kbp/min	
Ciclo 5:	72°C	x 10 min	

Tabella 3.2: Condizioni di reazione di PCR e programma del termociclatore.

3.2.4 MUTAGENESI SITO-SPECIFICA

La reazione di mutagenesi sito-specifica su IN è stata realizzata mediante il kit *QuickChange II XL Site-Directed Mutagenesis* (Stratagene). Questo kit risulta utile per creare mutazioni puntiformi, sostituire amminoacidi e inserire o eliminare singoli o multipli residui con un'efficienza di più dell'80%. Il materiale di partenza è un vettore a

doppio filamento di DNA contenente l'inserito di interesse. Il vettore viene poi replicato utilizzando due primer entrambi contenenti la mutazione di interesse e complementari ai due filamenti opposti del plasmide. Dopo l'appaiamento di questi oligonucleotidi la DNA polimerasi estende i primer generando il plasmide mutato. Il sistema prevede quindi la digestione del DNA parentale mediante l'endonucleasi DpnI che è specifica per il DNA metilato. Il vettore con la mutazione desiderata viene poi trasformato in cellule ultracompetenti XL10-Gold fornite nel kit.

	Mix di reazione
Buffer	5µl Buffer 10x
dNTPs	1 µl
Primer 1	125 ng
Primer 2	125 ng
DNA templato	10 ng
DNA polimerasi PfuUltra HF	1 µl
Quik Solution	3 µl
H2O	Fino a un volume di 50 µl

PROGRAMMA:

Ciclo 1:	95°C	x 1 min	
Ciclo 2:	95°C	x 50 sec	} 18 cicli
Ciclo 3:	60°C	x 50 sec	
Ciclo 4:	68°C	x 1 min/kb plasmide	
Ciclo 5:	68°C	x 7 min	

Tabella 3.3: Condizioni di reazione di PCR e programma del termociclatore per la reazione di mutagenesi sito-specifica su IN.

3.2.4.1 MUTAGENESI INDOTTA MEDIANTE RADIAZIONI UV

La mutagenesi random dell'IN è stata indotta utilizzando i raggi UV. Abbiamo utilizzato come sorgente di UV è lo *Stratalinker[®]UV Crosslinker*. L'energia di esposizione è stata determinata sperimentalmente mediante rapporto “ $\mu\text{J}/\text{cm}^2$ di esposizione agli UV/colonie cresciute dopo la trasformazione batterica”, stabilendo che il valore ottimale di energia è di $5.000 \mu\text{J}/\text{cm}^2$ per 150 ng di DNA risospeso in 20 μl di acqua. I prodotti del trattamento con UV sono stati usati per trasformare batteri elettrocompetenti.

3.2.5 PURIFICAZIONE DI FRAMMENTI DI DNA

I prodotti di PCR e i frammenti genici di digestione enzimatica sono stati purificati da gel di agarosio o mediante l'uso di colonnine filtranti. A tal fine sono stati seguiti le indicazioni fornite dalle ditte produttrici dei seguenti kit:

- *Kit QIAquick R pCR purification Kit Qiagen* per la purificazione di prodotti di PCR e per i prodotti di digestioni;
- *Zymoclean TM Gel DNA Recovery kit della Zymo research* per la purificazione dei prodotti da gel di agarosio;

L'integrità e la concentrazione del DNA purificato sono state verificate mediante elettroforesi su gel di agarosio

3.2.6 ELETTROFORESI SU GEL DI AGAROSIO E SCREENING DEI TRASFORMANTI

L'analisi e la purificazione dei vari campioni di DNA è stata ottenuta mediante corsa elettroforetica su gel di agarosio utilizzando come buffer di corsa TAE 0,5X. L'elettroforesi è stata effettuata applicando un voltaggio costante di 100V.

Lo screening dei cloni ottenuti è stato effettuato mediante *colony-PCR* e digestione enzimatica scegliendo:

- Un enzima che linearizza il vettore, digerendo all'interno dell'inserito clonato
- Un enzima che taglia due volte all'interno del plasmide, di cui una all'interno dell'inserito clonato in modo da generare due frammenti di lunghezza attesa.

Ulteriore metodo di screening utilizzato per verificare la direzionalità di inserti ligati in *blunt* è stato la PCR, al fine di verificare la direzionalità dell'inserito utilizzando dei primer interni all'inserito e dei primer esterni ad esso, e verificando se il prodotto di pcr aveva la lunghezza da noi attesa.

Un metodo veloce di screening per la positività di colonie cresciute su terreno LB-agar è la *colony-PCR*, metodica che consente di eseguire PCR di screening utilizzando come templat pochi microlitri di una soluzione ottenuta stemperando le colonie batteriche d'interesse in acqua e trattate a 95°C per 10 minuti e successivamente incubate in ghiaccio.

3.2.7 LIGAZIONE

I costrutti ottenuti in questo lavoro sono stati prodotti utilizzando l'enzima T4 DNA Ligasi (promega). A seconda delle estremità sticky o blunt del vettore e dell'inserito di

partenza la miscela di ligazione è stata incubata, seguendo le indicazioni del kit, a diverse tempistiche e temperature.

Il rapporto fra inserto e vettore usato è stato di 3:1 o 5:1 in base alle estremità sticky o blunt rispettivamente.

3.2.8 PROVE DI ESPRESSIONE

I vari costrutti di espressione ottenuti sono stati usati per trasformare il ceppo di *E. Coli* BL21 (DE3). Per ciascun costrutto abbiamo valutato l'espressione della chimera di almeno 5 trasformanti: ciascuna colonia è stata inoculata in 5 ml di LB con l'appropriato antibiotico selettivo e lasciata crescere O/N a 37°C in agitazione.

Il mattino seguente un'aliquota di 500 µl di ciascun inoculo è stata trasferita in 50 ml di terreno LB con antibiotico, e lasciata crescere a 37°C fino al raggiungimento di un OD₆₀₀ compreso tra 0,4 e 0,6 (crescita esponenziale).

A questo punto è stata indotta l'espressione dei transgeni aggiungendo l'induttore chimico, IPTG (*Isopropil-β-D-1-tiogalattopiranoside*) ad una concentrazione finale (Cf) di 0,5mM. Le colture batteriche sono state quindi lasciate a 37°C in agitazione. Sono state prelevate alcune aliquote dalle varie colture al tempo dell'induzione e a diversi tempi prestabiliti, (Tempo0, T2h, T3h e T5h). Per ciascuna aliquota è stata determinata la concentrazione della coltura batterica misurando allo spettrofotometro l'assorbanza a 600 nm (OD₆₀₀). Le aliquote sono state centrifugate a 6000 g per 5 minuti e il pellet batterico è stato conservato a -20°C.

3.2.9 SDS PAGE

I pellet batterici, raccolti durante le prove d'espressione, sono stati risospesi nel tampone di lisi denaturante *cracking buffer* 2x. I pellet sono stati risospesi con un volume di *cracking buffer* proporzionale agli OD₆₀₀ (ad esempio per OD 0,6 = 60 µl). I campioni sono stati bolliti a 95°C per 5 minuti per favorire la denaturazione delle proteine e sono stati corsi su gel di acrilamide al 10% mediante SDS-PAGE (*sodium Dodecyl Sulfate-Polyacrilamide Gel Electrophoresis*). La migrazione è avvenuta effettuando un voltaggio costante di 120 V per 2 h circa nel tampone tris-glicina.

Al termine della corsa elettroforetica i gel sono stati sottoposti a colorazione con *Coomassie Brilliant blue* per 1 h a RT; l'eccesso di colorante è stato poi rimosso trattando il gel per 20 minuti con la soluzione *Destaining* e infine in acqua distillata. Nel caso di analisi Western Blot, i gel migrati sono stati sottoposti direttamente a trasferimento su membrana evitando la colorazione con *Coomassie Brilliant blue*.

I gel di acrilamide e i diversi tamponi utilizzati sono stati preparati secondo le seguenti ricette:

Running Gel 10%	Per 10 ml	Stacking gel 5%	Per 3 ml
H2O	4 ml	H2O	2,1 ml
mix acrilamide 30%	3,3 ml	mix acrilamide 30%	0,5 ml
Tris 1,5 M pH 8.8	2,5 ml	Tris 1M pH 6.8	380 µl
SDS 10%	100 µl	SDS 10%	30 µl
Ammonio persolfato 10%	100 µl	Ammonio persolfato 10%	30 µl
Temed	4 µl	Temed	3 µl

Cracking buffer	Tris-glicina	Soluzione Coomassie	Soluzione Destaining
Blu Bromofenolo 0,02%	di Tris 25mM	EtOH 35%	MetOH 200 ml
Tris-HCl pH 6.8	0,09M Glicina pH8.3	MetOH 5%	Acido acetico 100 ml
Glicerolo 20%	SDS 0,1%	Acido acetico glaciale 10%	H2O distillata 700 ml
DTT 0,1%		Coomassie brillant blu 0,1%	
SDS 2%			

Tabella 3.4: Composizione dei gel di poliacrilamide e dei tamponi utilizzati.

3.2.10 WESTERN BLOT

In seguito a SDS-PAGE alcuni gel, quando è stato necessario, sono stati sottoposti direttamente ad analisi di Western Blot. La membrana (*Hydrabond-PVDF transfer membrane*, Amersham) è stata attivata con MetOH per 10 minuti in oscillazione ed è stata effettuata poi 3 lavaggi da 5 minuti in acqua. Dopo aver allestito la cassetta di trasferimento precedentemente equilibrata in tampone fosfato 1X, è stata inserita nella cella di trasferimento e le proteine sono state trasferite ad un amperaggio costante di 1200mA per due ore in ambiente refrigerato in tampone di trasferimento. Al termine del processo la membrana è stata colorata per 10 minuti con rosso *Ponceau*, per verificare l'effettivo trasferimento delle proteine sulla membrana, e decolorata completamente lavando con acqua distillata. In seguito la membrana è stata incubata con una soluzione di latte al 5% in PBS/Tween 0,05% per 1 ora in agitazione per saturare i siti di legame

aspecifici. Segue poi un lavaggio con PBS/Tween 0,05% per rimuovere l'eccesso di latte. La membrana è stata poi incubata, per 3 ore in agitazione, con una diluizione 1:1000 di un anticorpo primario monoclonale di topo anti IN (Santacruz Biotechnology) e, dopo 3 lavaggi di 5 minuti in PBS/Tween 0,05%, è stata incubata 1 ora in agitazione con una diluizione 1:1000 di anticorpo secondario di coniglio anti topo coniugato con HRP (Dako). Al termine dell'incubazione si effettuano 3 lavaggi da 5 minuti con PBS/Tween 0,05%, al termine dei quali si è aggiunta la soluzione di sviluppo, contenente il substrato per HRP. Tale soluzione è stata rimossa alla comparsa delle bande specifiche per l'integrasi, ed abbiamo bloccato la reazione lavando la membrana con acqua distillata.

Tampone fosfato di trasferimento 20X x 6Litri	Soluzioni di sviluppo
Na ₂ HPO ₄ H ₂ O 453,9 g	A: 100 ml TBS 1X + 60 µl H ₂ O ₂
NaH ₂ PO ₄ H ₂ O 71,3 g	B: 20 ml MetOH + 60 mg OrtoCloroNaftolo
NaN ₃ 1,2 g	

Tabella 3.5: Composizione dei tamponi utilizzati per il WB

3.2.11 PURIFICAZIONE DELLA CHIMERA

Una volta selezionati per ciascuna variante di chimera i trasformanti con un profilo di espressione della chimera ottimale abbiamo proceduto all'induzione dell'espressione della chimera su larga scala e alla relativa purificazione della proteina.

Per ciascuna variante abbiamo allestito una cultura pre-inoculo con un volume di 25ml, incubata O/N a 37°C in presenza di antibiotico ed in costante agitazione. Di questa cultura abbiamo prelevato 10ml ed inoculati in 500ml di terreno LB addizionato di

antibiotico, ed incubato a 37°C in costante agitazione fino al raggiungimento della fase esponenziale di crescita, corrispondente ad un OD₆₀₀ di 0,6 OD. Come in precedenza l'espressione della chimera è stata indotta aggiungendo alla coltura IPTG ad una concentrazione finale di 0,5mM, ed incubato a 37°C in costante agitazione per 4ore.

Al termine dell'incubazione la coltura batterica viene pellettata e sottoposta a 3 cicli di congelamento/scongelo a -80°C/4°C. Il pellet è risospeso in buffer 1 a cui abbiamo aggiunto Lisozima ad una Cf di 5 mg/ml e PMSF ad una Cf di 1mM.

Il pellet risospeso è omogeneizzato a bassa velocità in ghiaccio fino alla completa scomparsa della viscosità della miscela di lisi. Il lisato così ottenuto è stato incubato per 30 minuti a RT in agitazione orizzontale. Il lisato è stato poi trasferito in tubi da Ultra centrifuga e pellettato centrifugando a 13.250 rpm per 25 minuti a 20°C. Il surnatante ottenuto è stato prelevato e conservato a -80°C per analisi di SDS-PAGE. Il pellet è stato risospeso in buffer 2, contenente urea, un agente caotropico in grado di mantenere in soluzione le proteine insolubili, e omogeneizzato 2 volte a bassa velocità per 2 minuti.

Il lisato è stato nuovamente sottoposto a centrifugazione, alle medesime condizioni descritte in precedenza, ed il pellet formatosi è stato poi risospeso, in buffer 2.1 ed incubato per 30 minuti a RT in agitazione orizzontale. Nuovamente il lisato è stato sottoposto a centrifugazione a 13.000 rpm per 35 minuti a RT. Il surnatante è stato raccolto e conservato a -80°C. Il pellet è stato risospeso con PBS sterile e successivamente omogeneizzato a bassa velocità 2 volte per 2 minuti. Dopo un'ulteriore centrifuga a 13.000 rpm per 35 minuti a RT, è stato rimosso il surnatante e congelato a -80°C. Il pellet ottenuto, contenente i corpi di inclusione batterici purificati in cui è sequestrata la proteina è stato risospeso in buffer Lysis/Wash e la sospensione ottenuta è stata omogeneizzata per 2 minuti 5 volte a massima velocità. Con un'ulteriore centrifugazione a 13.000 rpm per 35 minuti a RT abbiamo isolato nuovamente il pellet e valutato la massa su bilancia di precisione. Il pellet è stato risospeso con un volume di buffer di solubilizzazione proporzionale alla massa del pellet (10 ml/g) ed incubato in agitazione orizzontale per 16 ore a 100 rpm a RT. Al termine dell'incubazione abbiamo isolato il surnatante centrifugando a RT per 35 minuti a 12.500 rpm; il surnatante così isolato è stato sottoposto a filtrazione attraverso filtri da 0.45 µm e successivamente da

0.22 μm .

L'intero purificato ottenuto è stato risolto mediante SDS-PAGE a 100 V. Al termine della corsa il gel è stato lavato in acqua distillata per eliminare residui di buffer ed immerso in una soluzione di cloruro di rame per 5 minuti a RT. Abbiamo exciso il frammento di gel corrispondente alla banda di peso molecolare desiderato e raccolto in un tubo da dialisi contenente tampone. Il tubo da dialisi contenente il frammento è stato sigillato con clamps, immerso in una cella elettroforetica orizzontale in tampone di elettroeluizione e sottoposto ad una corrente con voltaggio costante di 100V per 30 minuti, in modo favorire la migrazione delle proteine al di fuori del frammento di gel. Al termine dell'incubazione è stato invertita la polarità degli elettrodi ed è stata applicata una corrente costante di 100V per 30 secondi per permettere alle proteine adese alla superficie del tubo da dialisi, di immettersi nel liquido di elettroeluizione. La proteina elettroeluita è stata successivamente trasferita in un nuovo tubo da dialisi, sigillato con clumps, e sottoposta a dialisi in un litro di buffer di dialisi per 4-6 ore; al termine dell'incubazione il liquido di dialisi è stato sostituito con uno fresco e la dialisi è proceduta per ulteriori 16h. La proteina elettroeluita e dializzata è stata raccolta in sterilità e concentrata mediante colonne Centricon 50K (ditta) seguendo le indicazioni del manuale. L'eluato così ottenuto è stato aliquotato in aliquote da 10 μL e conservato a -80°C . Il concentrato rimasto sul filtro della colonna dopo centrifugazione, è stato diviso in due aliquote di cui una è diluita con glicerolo a Cf del 50%, mentre la seconda aliquota è portata a volume di 15 ml per poi ri-procedere con la concentrazione proteica.

I buffer utilizzati nel processo della solubilizzazione della proteina sono riportati nella tabella sottostante.

Buffer 1	Buffer 2	Buffer 2.1
Tris-HCl 20 mM pH 8	Tris-HCl 20 mM pH8	Tris-HCl 20 mM pH8
NaCl 500 mM	NaCl 500 mM	NaCl 500 mM
DTT 1 mM	Urea 4 M	Urea 10 M
DOC 2%	Triton 2%	Triton 2%
Triton 4.7%		

Buffer Lysis/Wash	Buffer Solubilization	Buffer di Dialisi
Na Phosphate Buffer 50 mM pH7	Na Phosphate Buffer 50 mM pH7	Na Phosphate Buffer 50 mM pH7
	SDS 1%	SDS 1%
		DTT 10 mM

Tabella 3.6: Composizione dei tamponi utilizzati nella solubilizzazione delle chimere

3.2.12 EMSA

Le proteine purificate sono state sottoposte a saggi di legame in presenza di DNA bersaglio per verificarne l'attività ed il corretto folding.

Sono stati incubati in 15 µl di volume finale 10 nM di DNA bersaglio e 200nM di proteina ricombinante in diversi tamponi contenenti Tris. Il tutto è stato incubato per 1h a 37°C. Al termine dell'incubazione la miscela di reazione è stata caricata su un gel di agarosio allo 0,8% e sottoposta alla corsa elettroforetica per 1h a 100V in 0,5X TAE buffer.

4. RISULTATI

4.1 STRATEGIA E COSTRUZIONE DELLA CHIMERA

L'idea alla base di questo progetto è di ottenere una proteina chimerica, costituita da un dominio DNA binding (IN) e da un dominio con attività nucleasica (*FokI*), capace di eradicare il provirus di HIV integrato nella cellula ospite.

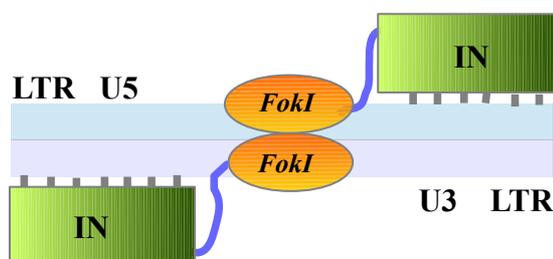


Figura 4.1: Rappresentazione grafica della chimera funzionale in cui i domini di IN sono legati agli LTR del DNA bersaglio e i domini di *FokI* sono dimerizzati.

La chimera WT è stata ottenuta amplificando mediante PCR le sequenze di IN, di L3 e di *FokI*, (i primer utilizzati per amplificare l'IN sono il Fw IN e il Rev IN, per il L3 è stato utilizzato la coppia di primer L3/RevL3 e per *FokI* la coppia Fw *FokI*/Rev *FokI*; le condizioni di PCR utilizzate sono descritte nella tabella 3.2), utilizzando come template costruiti, già presenti in laboratorio, contenenti rispettivamente le sequenze di IN, L3 e *FokI*.

L'avvenuta amplificazione è stata verificata mediante elettroforesi su gel di agarosio. Gli amplificati ottenuti sono quindi stati purificati (vedi paragrafo 3.2.5), sottoposti a digestione enzimatica e clonati all'interno del vettore di propagazione pUC19.

La chimera così ottenuta, è stata utilizzata come template di PCR per ottenere diverse chimere mutate:

Chimera D116A: il costrutto è stato ottenuto mediante l'introduzione della mutazione Asp116Ala nel dominio catalitico dell'Integrasi utilizzando il kit *QuickChange II XL Site-Directed Mutagenesis* (Stratagene) seguendo le indicazioni fornite dalla casa produttrice.

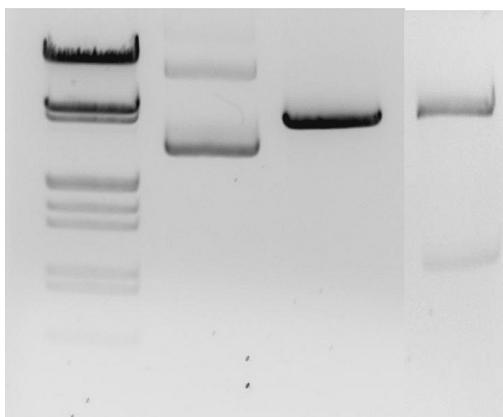


Figura 4.2: Analisi della chimera D116A mediante digestione enzimatica con Sall e SacII che tagliano rispettivamente all'interno del vettore pUC19 e all'interno del nostro inserto INL3FokI. Lane 1: Marker IV (Roche); lane 2: costrutto (chimera D116A) non digerito; lane 3: costrutto digerito con Sall; lane 4: costrutto digerito con Sall/SacII.

- **Chimere INDEL1, INDEL2 e INDEL3:** i costrutti presentano rispettivamente delezioni di 11, 18 e 27 aa nella regione C-terminale dell'IN. Tali delezioni sono state ottenute mediante reazioni di PCR, utilizzando i primer e le condizioni di reazione riportate nelle tabelle 3.1 e 3.2.

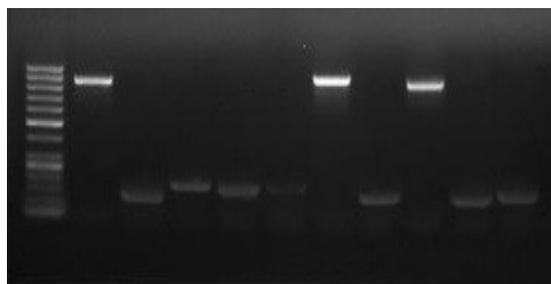


Figura 4.3: Screening delle chimere DEL1, DEL2 e DEL3 mediante PCR *colony*. Lane 1: Marker VIII; lane 2: DEL1; lane 7: DEL2; lane 9: DEL3.

Chimera D116A/delCTD: il costrutto presenta l'IN mancante del dominio CTD, ottenuto mediante reazione di PCR utilizzando i primer Fw IN e Rev IN-NoCTD (vedi tabella 3.1) e come template la chimera D116A.

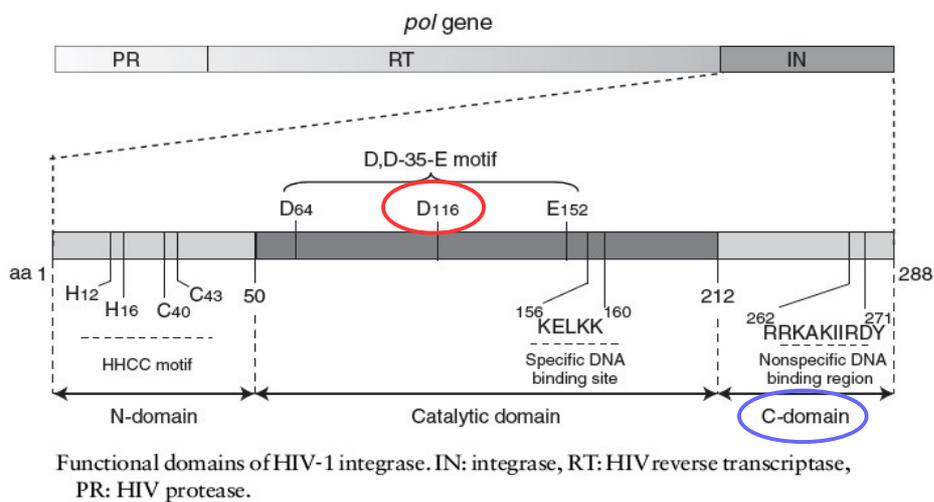


Figura 4.4: Rappresentazione grafica dei domini dell'IN in cui sono state effettuate le delezioni (in blu) e la mutazione puntiforme D116A (in rosso)

Queste mutazioni puntiformi e delezioni sono state create per attenuare o abolire l'attività catalitica dell'enzima e ridurre il legame aspecifico di IN al DNA, dovuto al dominio CTD dell'IN stessa. In tal modo ci aspettiamo di ottenere una chimera capace di riconoscere in modo selettivo le LTR di HIV ed essere in grado di tagliare il doppio filamento a valle del sito di riconoscimento.

4.2 STRATEGIA E COSTRUZIONE BERSAGLI

Parallelamente alla costruzione delle chimere, sono stati progettati e costruiti tre bersagli molecolari, necessari per lo studio dell'attività della proteina ricombinante.

I bersagli sono stati progettati per mimare le giunzioni LTR-LTR trovate nelle forme circolari di DNA virale. Essi sono costituiti dal plasmide pCR8/GW/TOPO recante:

- gli elementi LTR U3 e U5, rappresentanti i substrati candidati che la nostra chimera dovrebbe riconoscere e legare al fine di permettere l'attività idrolitica di FokI.
- la sequenza del gene *ccdB*, ricavato dal sistema *ccd* del plasmide F di *E.Coli*, che codifica la proteina tossica *ccdB* in grado di interferire con l'attività della DNA girasi.
- il promotore trascrizionale dell'operone *araBAD* (arabinosio), che controlla la trascrizione del gene tossico mediante induzione di arabinosio (figura 4.5 A).

In laboratorio sono stati progettati tre bersagli che differiscono fra loro per la presenza di una sequenza spaziatrice di 0, 20 o 40 bp, interposta tra gli elementi LTR U3 e U5 (figura 4.5 C). L'utilizzo delle sequenze spaziatrici potrebbe essere necessario per favorire, da un punto di vista sterico, la dimerizzazione funzionale di due molecole chimeriche tale da attivare la funzione nucleasica del complesso.

E' stato creato inoltre un costrutto costituito dal plasmide pACYC184 e il gene LacY (figura 4.5 B) codificante per la β -galattosio permeasi. Tale enzima permette l'ingresso dello zucchero (nel nostro caso l'arabinosio) all'interno della cellula che, a sua volta, attiva il promotore pBAD inducendo la trascrizione del gene tossico.

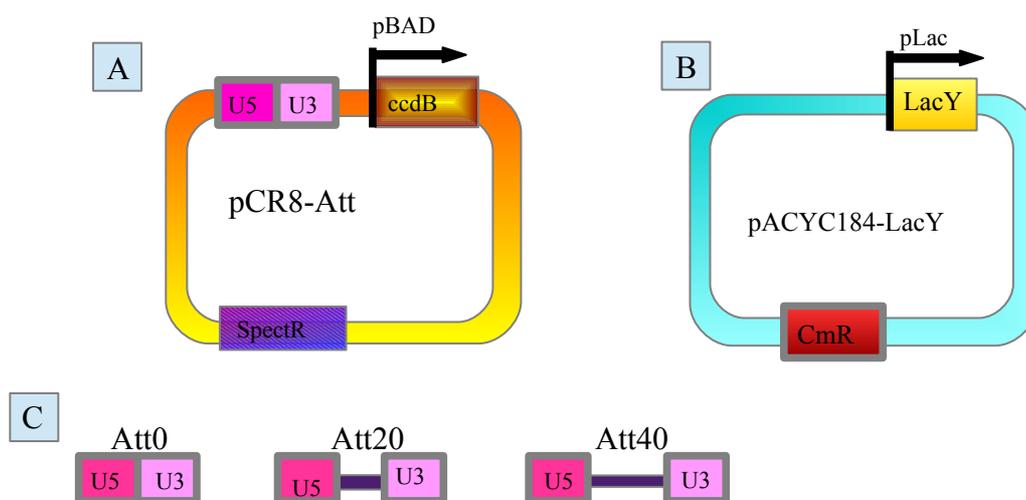


Figura 4.5: In A è rappresentato il modello grafico del plasmide bersaglio, in B il plasmide recante il gene LacY e in C le sequenze spaziatriche interposte tra gli LTR.

4.3 ESPRESSIONE E PURIFICAZIONE DELLA CHIMERA

Le chimere delCTD e DEL2 sono state espresse, mediante induzione con IPTG, nelle cellule batteriche BL21 (DE3) (vedi parag. 3.2.8)

I vari campioni raccolti ai diversi tempi sperimentali sono stati corsi su gel di acrilamide al 10% e sottoposti al trasferimento delle proteine su membrana mediante Western blotting. L'incubazione della membrana con l'anticorpo monoclonale murino anti-Integrasi, ha rivelato che l'espressione ottimale della proteina è stata ottenuta a 37°C, 0,5 mM di IPTG e a 4h dall'induzione.

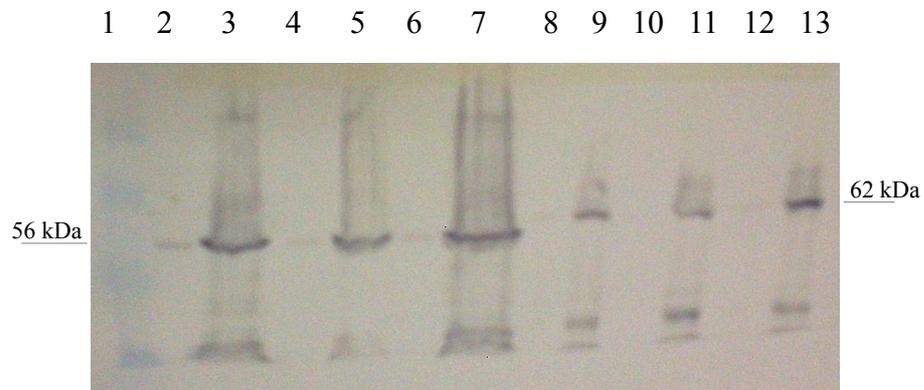


Figura 4.6: WB anti IN dell'espressione della chimera delCTD (lanes 2-7) e della chimera DEL2 (lanes 8-13). Lane 1: Prestained Protein Molecular Weight Marker; lanes 2-3, 8-9: espressione della chimera indotta da IPTG a T0 e T5h a RT; lanes 4-5, 10-11: espressione della chimera indotta da IPTG al T0 e T24h a 4°C; Lanes 6-7, 12-13: espressione della chimera indotta da IPTG al T0 e T4h a 37°C.

Prima di procedere con le varie fasi della purificazione è stata verificata la solubilità della proteina, raccolta in varie fasi della lisi batterica, mediante corsa elettroforetica su gel di acrilamide al 10% e analisi Western Blot anti IN (figura 4.7).

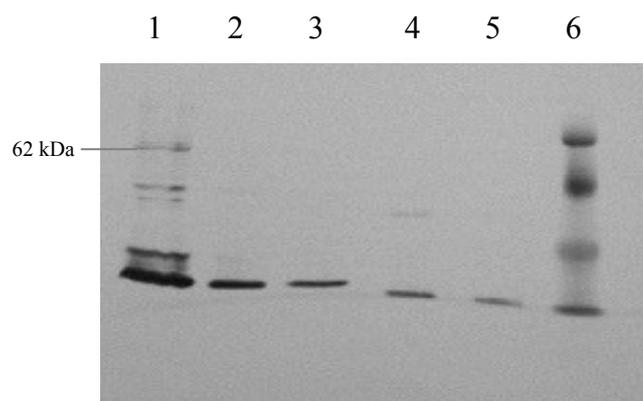


Figura 4.7: Verifica della solubilità dell'IN mediante Western Blot. Lane 1: Frazione totale; lanes 2-5: supernatanti ottenuti dalla lisi batterica; lane 6: Prestained Protein Molecular Weight Marker.

Dalla figura 4.7 si può osservare chiaramente che la nostra proteina ricombinante non è presente nella frazione solubile, ma solamente nella frazione totale (lane 1). Abbiamo quindi ipotizzato che la cellula batterica sequestri l'enzima all'interno di corpi di inclusione. Dopo diversi tentativi di purificazione è stato messo a punto il protocollo di purificazione (vedi par. 3.2.11) che ha permesso di ottenere la proteina (figura 4.8).

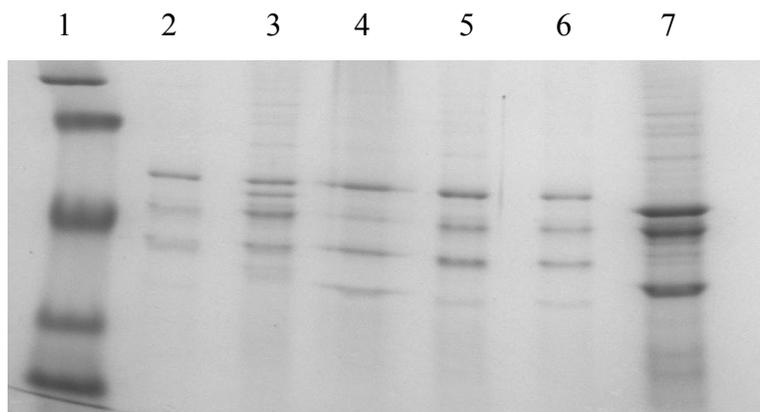


Figura 4.8: WB anti IN dopo purificazione della proteina. Lane1: Prestained Protein Molecular Weight Marker; lane2: proteinaWT; lane3: proteinaD116A; lanes4-6: proteinaDel1, Del2 e Del3; Lane7: proteina D116A/dCTD.

4.4 EMSA

Per determinare l'affinità di legame della chimera per il DNA bersaglio, sono stati condotti dei saggi EMSA (*Electrophoretic Mobility Shift Assay*; vedi par. 3.2.12). La modalità di esecuzione di questo saggio consiste nell'incubare la proteina chimerica con il DNA bersaglio in modo tale da analizzare l'attività di legame al DNA. I complessi DNA-proteina vengono quindi separati dal DNA non legato mediante elettroforesi in un gel allo 0,8% di agarosio. La proteina ritarda la corsa del frammento di DNA a cui si lega, quindi il DNA libero migra più velocemente attraverso il gel rispetto ai complessi DNA-proteina.

Gli esperimenti sono stati condotti con una quantità costante di DNA bersaglio

(pCR8Att0 e pCR8Att20) e di proteina (chimera WT, chimera DEL2) in tamponi salini contenenti ioni zinco 50 μ M.

Il controllo negativo della reazione è rappresentato dall'incubazione dell'enzima in presenza del plasmide pCR8 privo degli LTR, e dall'incubazione di una proteina, la β -galattosidasi (purificata nelle medesime condizioni della chimera), con il plasmide pCR8Att0 linearizzato.

Dai risultati dei saggi EMSA (figura 4.9) si vede che la chimera DEL2 forma un complesso con il DNA bersaglio, infatti presenta ridotta mobilità elettroforetica se paragonata alla proteina incubata in presenza di DNA bersaglio senza LTR (controllo negativo) e della proteina β -galattosidasi in presenza del DNA bersaglio.

Inoltre si può notare che la proteina DEL2 lega, seppur in maniera minore, il controllo negativo. Da quest'analisi si deduce che l'incubazione Chimera-DNA bersaglio non ha provocato clivaggio rilevabile del DNA bersaglio.

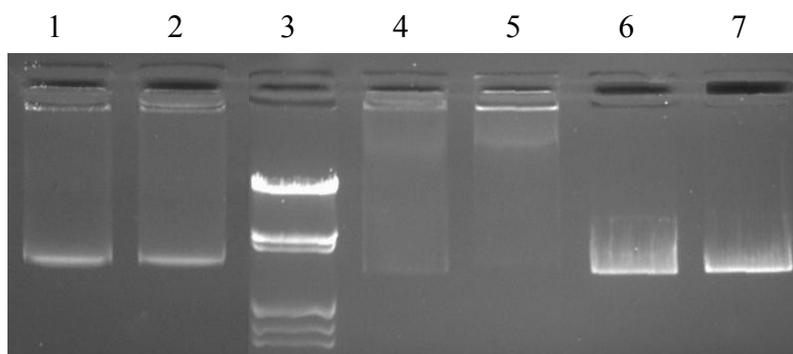


Figura4.9: Lane 1: Chimera INwt + pCR8ccdB; lane 2: Chimera INwt + pCR8-Att0; lane 3: Marker III; lane 4: Chimera INDEL2 + pCR8ccdB; lane 5: Chimera INDEL2 + pCR8-Att0; lane 6: β -galattosidasi + pCR8ccdB; lane 7: β -galattosidasi + pCR8-Att0.

Per aumentare la specificità di legame proteina-DNA bersaglio e permettere a FokI di svolgere la sua azione, sono state ottimizzate le condizioni di reazione testando l'attività delle chimere in diversi tamponi e variando la concentrazione di ioni zinco. I saggi EMSA sono stati quindi svolti nei vari tamponi NEB della ditta *Biolabs*, e sottoposti a corsa elettroforetica su gel di agarosio. L'analisi ha mostrato un incremento dell'attività

di legame in presenza di 50 mM di NaCl e a pH di 7,9 (tampone NEB2), infatti il complesso proteina-DNA bersaglio presenta una ridotta mobilità elettroforetica (figura 4.10).

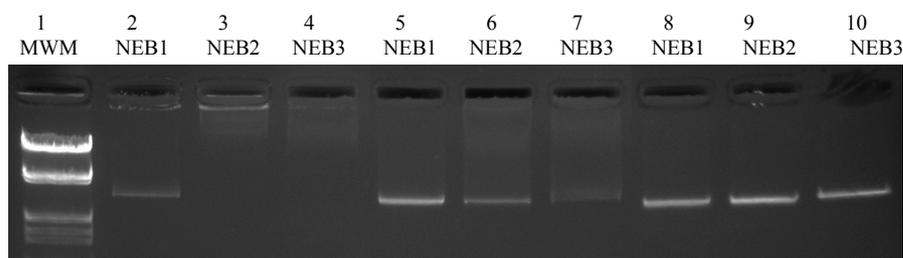


Figura 4.10: Lane 1: Marker III; lanes 2-4: chimera DEL2+pCR8Att20; lanes 5-7: chimera DEL2+pCR8ccdB; lanes 8-10: pCR8Att20 senza la chimera.

La diminuzione o l'abolizione della concentrazione di ioni zinco nella reazione di legame, ha influito negativamente sulla reazione, in quanto la chimera ha perso la capacità di formare complessi con il DNA bersaglio (figura 4.11). Questo risultato ha confermato ulteriormente l'importanza dello ione zinco nella reazione che, oltre a favorire la formazione di dimeri e tetrameri di IN, è necessario per incrementare l'attività di IN Mg^{2+} dipendente.



Figura 4.11: Saggi EMSA a diverse concentrazioni di ioni zinco.

E' stata inoltre testata la capacità della chimera WT di legare un oligonucleotide recante gli LTR di 90 bp e un oligonucleotide di 130bp (controllo negativo). Le condizioni del saggio EMSA sono identiche a quelle sopra indicate. I risultati purtroppo sono stati deludenti in quanto la chimera WT non si lega a nessuno dei due frammenti (figura 4.12).

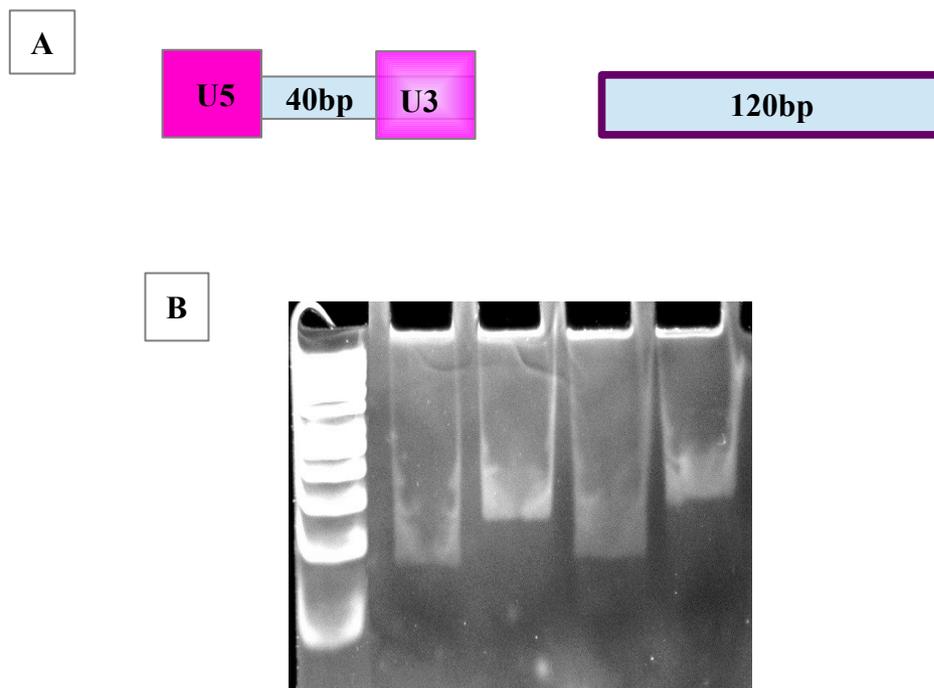


Figura 4.12: In A sono mostrate la sequenza LTR+40bp e l'oligonucleotide 120bp . Figura B: lane1 Marker; lane2 LTR+40bp; lane3: oligonucleotide 120bp; lane 4: LTR+40bp + chimeraWT; lane5: oligonucleotide 120bp+ chimeraWT.

Poiché i risultati ottenuti in vitro non sono stati soddisfacenti, in quanto in nessuna delle condizioni di reazione a cui la chimera è stata sottoposta si è ottenuto il clivaggio del DNA bersaglio ma solo un'attività di legame, è stato ideato e creato il sistema della selezione in vivo in cellule procariotiche.

4.5 SISTEMA DI SELEZIONE IN VIVO

Il sistema di evoluzione in vivo è basato sulla discriminazione di mutanti positivi e sulla loro continua mutagenizzazione e selezione, allo scopo di ottenere chimere con mutazioni funzionali.

Per ottenere ciò, è stato ideato un sistema di selezione in cui le cellule batteriche, esprimenti una chimera non funzionale, non sopravvivano. Le cellule batteriche utilizzate in questo sistema sono BW25.141, un ceppo ingegnerizzato di *E. Coli* capace di degradare rapidamente il DNA lineare. Questi batteri sono stati trasformati con il plasmide pCR8 recante gli LTR e il gene tossico *ccdB*, e con il plasmide pACYC184 recante il gene per LacY. Il promotore pBAD, posto a monte del gene *ccdB*, permette la trascrizione del gene tossico solo in presenza di arabinosio, quindi se trascritto induce morte cellulare.

Nel caso in cui il batterio produca una chimera IN-FokI funzionale, che permetta il taglio del plasmide contenente *ccdB*, esso sarà in grado di crescere anche in presenza dell'induttore del gene tossico, arabinosio (figura 4.13).

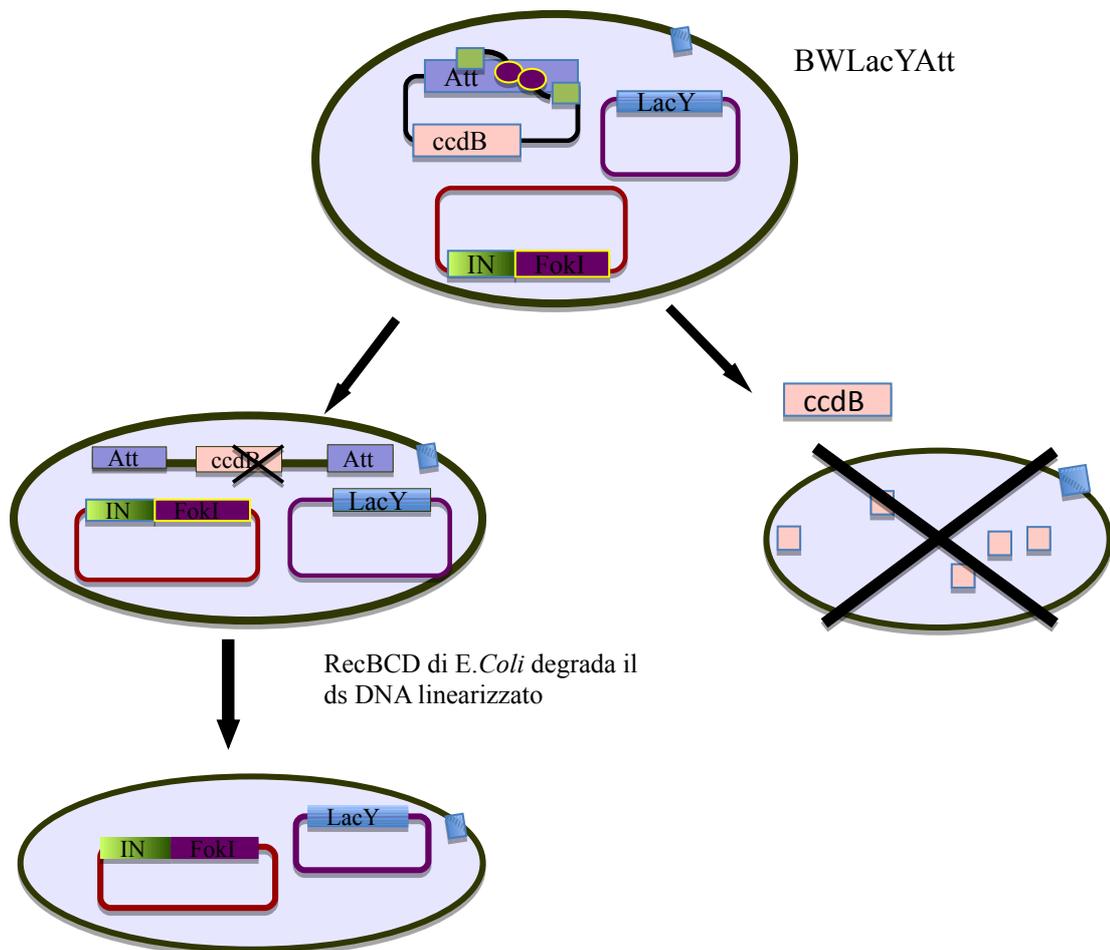


Figura 4.13: Illustrazione grafica della selezione in vivo. Il sistema di selezione collega l'evento di idrolisi operato dalla chimera alla sopravvivenza cellulare.

Per ogni ciclo di selezione sono stati utilizzati 50 ng di DNA mutato. In tal modo si ha la probabilità di ottenere un numero di mutazioni favorevoli tra le 50 e le 500, probabilità calcolata ipotizzando che, nelle condizioni sperimentali da noi utilizzate, è possibile ottenere almeno una mutazione vantaggiosa ogni 1.000, su 10^6 batteri trasformati per ug di DNA.

L'intero costrutto pUC19 + D116A/delCTD è stato sottoposto a mutagenesi random UV, per cui si potranno ottenere i seguenti risultati:

- Se le mutazioni sono favorevoli, cioè aumentano la specificità dell'IN nel riconoscere il DNA bersaglio, attribuiranno alla chimera una migliore capacità di riconoscere il DNA e permettere a *FokI* di idrolizzarlo, con conseguente crescita batterica.
- Se invece le mutazioni inattivano il gene della resistenza antibiotica, si avrà morte batterica a causa della presenza dello stesso antibiotico nelle fasi di selezione; lo stesso risultato si otterrà se viene mutata l'origine di replicazione.
- Se le mutazioni inattivano il promotore che regola l'espressione della chimera, si avrà morte batterica in quanto non verrà più controllata la trascrizione del gene di suicidio *ccdB* espresso dai batteri BW25.141.

La chimera D116A/delCTD, sottoposta a mutagenesi UV (vedi par. 3.2.4.1), è stata utilizzata per trasformare il ceppo Top10 di *E.Coli*. Dai trasformanti ottenuti è stato estratto il DNA plasmidico mediante miniprep, in maniera tale da ottenere una libreria espressiva di DNA mutagenizzato. Lo step successivo ha previsto la trasformazione di batteri BW25.141 doppi trasformanti, cioè aventi il plasmide pCR8 + *ccdB* + LTR (*Att0*, *Att20* e *Att40*) e il plasmide pACYC184 + *LacY*, con il DNA mutagenizzato. Come controllo negativo è stato utilizzato il ceppo BW25.141*ccdBLacY* privo di LTR.

Rispetto alle normali trasformazioni batteriche, nella fase post-elettroporazione è stato aggiunto l'induttore per l'espressione della chimera al brodo di crescita batterica, e lasciato ad incubare 1h a 37°C, in modo da permettere alla chimera, espressa in quantità sufficiente, di competere efficacemente con il *ccdB* per la sopravvivenza del batterio.

Al fine di selezionare le colonie prive del bersaglio LTR, sono stati seminati 50 ul di coltura batterica in piastre LB+agar contenenti:

- gli antibiotici ampicillina e cloramfenicolo, e lo zucchero arabinosio (ara +)
- gli antibiotici ampicillina e cloramfenicolo, senza arabinosio (ara -)

Le piastre sono state incubate per 16 h a 37°C, e successivamente le ara+ sono state

lasciate ad incubare altri 2 giorni a 25°C, in quanto dopo 16h le colonie erano ancora di dimensioni piccole.

Il numero di trasformanti ottenuti risulta essere maggiore nel ceppoAtt0 (150 colonie) rispetto ad Att20 (83 colonie) e Att40 (23 colonie), ciò fa presupporre che l'assenza della sequenza spaziatrice tra gli LTR faciliti la dimerizzazione dei domini catalitici di FokI della nostra chimera che in tal modo riesce ad idrolizzare il bersaglio. La crescita batterica in BW25.141ccdBLacY privo di LTR è stata nulla nelle piastre contenenti arabinosio, indicando che la chimera in assenza dei bersagli LTR non riesce ad idrolizzare il plasmide.

Per verificare la presenza effettiva della nostra chimera all'interno delle colonie batteriche, è stato estratto il DNA plasmidico da alcuni trasformanti ottenuti e verificata, mediante PCR colony, la presenza della sequenza dell'IN (i primer utilizzati sono: Fw IN116A e RevINnoCTD; per il programma di amplificazione si veda tabella 3.2). L'altezza del amplificato, 636 bp, corrisponde alla grandezza dell'IN deleta del dominio CTD (figura 4.14).

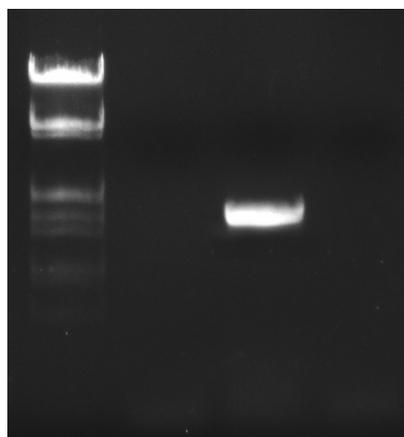


Figura 4.14: Screening della PCR colony mediante elettroforesi su gel di agarosio allo 0,8%. Lane 1: marker III; lane 3: prodotto di PCR.

Il DNA plasmidico, estratto dalle colonie ottenute dal primo ciclo di mutagenesi, è stato

utilizzato per ottenere la seconda generazione di mutanti. Abbiamo proseguito in questa linea per avere diverse generazioni di mutanti.

Dopo ogni ciclo di mutagenesi, è stato estratto il DNA plasmidico da alcune colonie scelte casualmente, allo scopo di sequenziarlo e confrontarlo con la sequenza della chimera non sottoposta a mutagenesi UV (tabella 4.1).

Dalle analisi risulta che negli Att0, dopo il primo ciclo di mutagenesi, il tasso di mutazione è pari allo 1,07% a livello della sequenza nucleotidica e di 1,38% a livello della sequenza amminoacidica rispetto alla chimera non mutagenizzata. Negli Att40 il tasso di mutazione a livello nucleotidico è di 1,88% e a livello amminoacidico è di 4,76%.

Negli Att20 il tasso di mutazione a livello nucleotidico è pari a 1,54% mentre a livello amminoacidico è pari a 3,70% nella terza generazione.

Da quest'analisi è stato visto, quindi, che alcune mutazioni insorte nel primo ciclo di mutagenesi restano invariate nei cicli successivi e in contemporanea sono comparse nuove mutazioni nelle generazioni successive. Queste mutazioni hanno portato alla trascrizione di chimere con amminoacidi differenti (tabella 4.3).

HIV gcatggacaagtagactgtagtccaggaatatggcaactagattgtacacatttagaagg
Chimera gcatggacaagtagactgtagtccaggaatatggcaactagattgtacacatttagaagg
1r.Att0UV.17.M1 gcgtggacaagtagactgtagtccaggaatatggcaactagattgtacacatttagaagg
3r.ATT20.2.E39_ gcgtaggacaagtagactgtagtccaggaatatggcaactagattgtacacatttagaagg
 ●***

HIV aaaagttatcctggtagcagttcatgtagccagtggatataatagaagcagaagttattcc
Chimera aaaagttatcctggtagcagttcatgtagccagtggatataatagaagcagaagttattcc
1r.Att0UV.17.M1 aaaagttatcctggtagcagttcatgtagccagtggatataatagaagcagaagttattcc
3r.ATT20.2.E39_ aaaagttatcctggtagcagttcatgtatccagtggatataatagaatcagaagttattcc
 *****●*****●*****

HIV agcagaaacagggcaggaaacagcatactttctcttaaaattagcaggaagatggccagt
Chimera agcagaaacagggcaggaaacagcatactttctcttaaaattagcaggaagatggccagt
1r.Att0UV.17.M1 agcagaaacagggcaggaaacagcatactttctcttaaaattagcaggaagacggccagt
3r.ATT20.2.E39_ agcagaaacagggcaggaaacagcatactttctcttaaaattagcaggaagacggccagt
 *****●*****

HIV aaaaacaatacatagagacaatggcagcaatttcaccagttactacgggtaaggccgcctg
Chimera aaaaacaatacatagagacaatggcagcaatttcaccagttactacgggtaaggccgcctg
1r.Att0UV.17.M1 aaaaacaatacatagagacaatggcagcaatttcaccagttactacgggtaaggccgcctg
3r.ATT20.2.E39_ aaaaacaatacatagagacaatggcagcaatttcaccagttactacgggtaaggccgcctg
 *****●●*****

HIV ttggtgggcggggaatcaagcaggaatttggatttcctacaatccccaaagtcaaggagt
Chimera ttggtgggcggggatcaagcaggaatttggatttcctacaatccccaaagtcaaggagt
1r.Att0UV.17.M1 ttggtgggcggggatcaagcaggaatttggatttcctacaatccccaaagtcaaggagt
3r.ATT20.2.E39_ ttggtgggcggggatcaagcaggaatttggatttcctacaatccccaaagtcaaggagt
 *****●*****●*****

HIV agtagaatctatgaataaagaattaaagaaaattataggacaggtgaagagatcaggctga
Chimera agtagaatctatgaataaagaattaaagaaaattataggacaggtgaagagatcaggctga
1r.Att0UV.17.M1 agtagaatctatgaataaagaattaaagaaaattataggacaggtgaagagatcaggccga
3r.ATT20.2.E39_ agtagaatctatgaataaagaattaaagaaaattataggacaggtgaagagatcaggccga
 *****●*****

HIV acatcttaagacagcagtagcaaatggcagtagttcatccacaattttaaaagaaaaggggg
Chimera acatcttaagacagcagtagcaaatggcagtagttcatccacaattttaaaagaaaaggggg
1r.Att0UV.17.M1 acatcttaagacagcagtagcaaatggcagtagttcatccacaattttaaaagaaaaggggg
3r.ATT20.2.E39_ acatcttaagacagcagtagcaaatggcagtagttcatccacaattataaaaagaaaaggggg
 *****●*****●*****

HIV gattggggggtacagtgcaggggaaagaatagtagacataatagcaacagacatacaaac
Chimera gattggggggtacagtgcaggggaaagaatagtagacataatagcaacagacatacaaac
1r.Att0UV.17.M1 gattggggggtacagtgcaggggaaagaatagtagacataatagcaacagacatacaaac
3r.ATT20.2.E39_ gattggggggtacagtgcaggggaaagaatagtagacattatagcaacagacatacctac
 *****●*****●*****●●*****

HIV taaagaattacaaaaaaaattacaaaaattcaaaattttcgggtttattacagggacag
Chimera taaagaattacaaaaaaaattacaaaaattcaaaattttcgggtttattacagggacag
1r.Att0UV.17.M1 taaagaattacaaaaaaaattacaaaaattcaaaattttcgggtttattacagggacag
3r.ATT20.2.E39_ taaagaattacaacaacaaa-----
 *****●*****

HIV cagagatccactttggaaaggaccagcaagctcctctggaagggtgaaggggcagtagt
Chimera cagaaatccactttggaaaggaccagcaagctcctctggaagggtgaaggggcagtagt
1r.Att0UV.17.M1 cagaaatccactttggaaaggaccagcaagctcctctggaagggtgaaggggcagtagt
3r.ATT20.2.E39_ cagaaatccactttggaaaggaccagcaagctcctctggaagggtgaaggggcagtagt

HIV aatacaagataatagtgacataaaaagtagtgccaagaagaaaagcaaagatcattagggga
Chimera aatacaagataatagtgacataaaaagtagtgccaagaagaaaagcaaagatcattagggga
1r.Att0UV.17.M1 aatacaagataatagtgacataaaaagtagtgccaagaagaaaagcaaagatcattagggga
3r.ATT20.2.E39_ -----

HIV ttatggaaaacagatggcaggtgatgattgtgtggcaagtagacaggatgaggattagcc
Chimera ttatggaaaacagatggcaggtgatgattgtgtggcaagtagacaggatgaggat---cc
1r.Att0UV.17.M1 ttatggaaaacagatgtcaggtgatgattgtgtggcaagtagacaggatgaggat---cc
3r.ATT20.2.E39_ -----

HIV gcggggaggaggaggaaagcgggtggaggaggcagcggcggcggcggtagcgcgatgcaagca
Chimera gcggggaggaggaggaaagcgggtggaggaggcagcggcggcggcggtagcgcgatgcaagca
1r.Att0UV.17.M1 gcggggaggaggaggaaagcgggtggaggaggcagcggcggcggcggtagcgcgatgcaagca
3r.ATT20.2.E39_ -----

HIV actagtcaaaagtgaactggaggagaagaaatctgaacttcgacataaattgaaatagtg
Chimera actagtcaaaagtgaactggaggagaagaaatctgaacttcgacataaattgaaatagtg
1r.Att0UV.17.M1 actagtcaaaagtgaactggag--agagaaatctgaacttcgacataaattgaaatagtg
3r.ATT20.2.E39_ -----

HIV gcctcatgaatataatgaattaattgaaattgccagaaattccactcaggatagaattct
Chimera gcctcatgaatataatgaattaattgaaattgccagaaattccactcaggatagaattct
1r.Att0UV.17.M1 gcctcatg-atataattgatt-----attgaatgcagaatcactcaga
3r.ATT20.2.E39_ -----

HIV tgaaatgaaggtaatggaatTTTTTatgaaagtttatggatatagaggtaaacatttggg
Chimera tgaaatgaaggtaatggaatTTTTTatgaaagtttatggatatagaggtaaacatttggg
1r.Att0UV.17.M1 tagaatctgaatgagtagatgattTTTTatg-aagtttatggatatagaggtaaacatttggg-
3r.ATT20.2.E39_ -----

HIV tggatcaaggaaaccggacggagcaatttatactgtcggatctcctattgattacggtgt
Chimera tggatcaaggaaaccggacggagcaatttatactgtcggatctcctattgattacggtgt
1r.Att0UV.17.M1 -----tgatcagaacgaaaccggagccatttatactgtcggatctcctattgat-----
3r.ATT20.2.E39_ -----

HIV gatcgtggataactaaagcttatagcggaggttataatctgccaatggccaagcagatga
Chimera gatcgtggataactaaagcttatagcggaggttataatctgccaatggccaagcagatga
1r.Att0UV.17.M1 -----
3r.ATT20.2.E39_ -----

HIV aatgcaacgatatgtcgaagaaaatcaaacacgaaacaaacatatcaaccctaataatgaatg
Chimera aatgcaacgatatgtcgaagaaaatcaaacacgaaacaaacatatcaaccctaataatgaatg
1r.Att0UV.17.M1 -----acg
3r.ATT20.2.E39_ -----

HIV gtggaaagtctatccatcttctgtaacggaatttaagttttatTTTgtgagtggtcactt
Chimera gtggaaagtctatccatcttctgtaacggaatttaagttttatTTTgtgagtggtcactt
1r.Att0UV.17.M1 gtgtggagtc-----ctgcatctaagctat-----
3r.ATT20.2.E39_ -----

```

HIV          taaaggaaactacaaagctcagcttacacgattaaatcatatcactaattgtaatggagc
Chimera      taaaggaaactacaaagctcagcttacacgattaaatcatatcactaattgtaatggagc
1r.Att0UV.17.M1 -----ccgagatttatctgcaactgcagccgag-
3r.ATT20.2.E39_ -----

HIV          tgttcttagtgtagaagagcttttaattggtggagaaatgattaaagccggcacattaac
Chimera      tgttcttagtgtagaagagcttttaattggtggagaaatgattaaagccggcacattaac
1r.Att0UV.17.M1 -----
3r.ATT20.2.E39_ -----

HIV          cttagaggaagtgagacggaaatttaataacggcgagataaaacttttaa
Chimera      cttagaggaagtgagacggaaatttaataacggcgagataaaacttttaa
1r.Att0UV.17.M1 -----agatattg
3r.ATT20.2.E39_ -----a

```

Tabella 4.2: Allineamento delle sequenze nucleotidiche di: Integrase di HIV-1 (in nero), chimera D116A/delCTD (in verde), Chimera 1° ciclo di mutagenesi (in blu) e chimera dopo 3° ciclo di mutagenesi (in rosso) .

5. DISCUSSIONE

L'AIDS è la sindrome provocata dall'infezione da HIV, attualmente trattabile con una terapia farmacologica a base di numerose combinazioni di farmaci antiretrovirali, ma ancora non guaribile. Ad oggi la terapia è l'HAART, che però è sensibile allo sviluppo delle resistenze e non è in grado di eradicare il virus dai pazienti infetti; per questo si stanno sviluppando una serie di approcci alternativi mirati a prevenire l'infezione, migliorare la terapia HAART o, come nel caso della tecnologia ZFN, eradicare il virus dalla cellula. Il nostro progetto appartiene a questo nuovo campo di ricerca.

L'obiettivo che ci siamo preposti è ottenere una proteina chimerica capace di riconoscere e recidere il provirus di HIV integrato nel genoma della cellula ospite. A tal fine sono state create diverse chimere, cioè Chimera WT, chimera D116A, chimera DEL1, chimeraDEL2, chimera DEL3 e chimera D116A/delCTD. Tutte costituite dalla sequenza dell'Integrasi di HIV, da una sequenza spaziatrice (L3) e dalla sequenza del dominio catalitico di FokI.

Abbiamo testato l'attività delle chimere D116A/delCTD e DEL2, facendole interagire con i bersagli molecolari contenenti la sequenza LTR. Dalle analisi, ottenute mediante saggi EMSA, abbiamo verificato che la chimera DEL2 e in parte la chimera D116A/delCTD presentano maggiore affinità di legame al DNA bersaglio rispetto alla chimera WT. Questi risultati hanno confermato l'ipotesi da noi fatta, vale a dire che opportune modifiche del dominio CTD del IN possono ridurre il legame aspecifico del IN al substrato. Ad oggi tuttavia non siamo riusciti ad ottenere in vitro l'attività completa della chimera, ma siamo riusciti ad ottenere solo un'attività di DNA-binding mentre l'attività di idrolisi non è stata rilevata.

Il sistema di selezione in vivo della chimera è nato quindi dall'idea di implementare un metodo di selezione positiva che sfrutta batteri opportunamente ingegnerizzati. La

chimera sottoposta a mutagenesi UV è stata trasformata nel ceppo BW25.141 recante il plasmide pCR8 + LTR + ccdB e il plasmide pACYC184 + LacY. La sopravvivenza delle cellule batteriche è legata alla funzionalità della chimera stessa, cioè se le mutazioni apportano alla chimera una migliore capacità di riconoscere ed idrolizzare il plasmide portante gli LTR e il gene tossico ccdB si avrà crescita cellulare al contrario in presenza di una chimera non funzionale si avrà l'espressione solo del gene tossico ccdB che causerà morte cellulare.

I risultati ottenuti ad oggi sono incoraggianti: dopo tre cicli di mutagenesi UV e selezione in vivo abbiamo ottenuto un numero di trasformanti maggiore nel ceppo BW25.141Att0, ciò potrebbe indicare che la mancanza della sequenza spaziatrice interposta tra gli LTR favorisce la dimerizzazione del dominio catalitico di FokI di due monomeri di chimera.

Dopo ogni ciclo, il DNA estratto da alcuni trasformanti BW25.141Att0 e BW25.141Att20 è stato sequenziato e confrontato con la sequenza della chimera non sottoposta a mutagenesi UV. Il tasso di mutazione risulta essere: di 1,07% a livello della sequenza nucleotidica mentre a livello della sequenza amminoacidica è di 3,0%.

Sulla base di questi risultati, per il futuro ci aspettiamo di:

- continuare con la metodica di mutagenesi UV in un sistema di selezione procariotico fino al decimo ciclo.
- Eseguire analisi statistiche al fine di calcolare il tasso di mutazione, indotto con la
- mutagenesi UV, così da comprendere quali sono le mutazioni utili per l'attività della chimera stessa.
- Analizzare l'attività di legame e di idrolisi della chimera mutagenizzata in vitro.
- Analizzare l'attività delle chimere in cellule eucariotiche.

6. BIBLIOGRAFIA

- 1 - Campbell, G.R., et al., *The glutamine -rich region of the HIV-1 Tat protein is involved in T-cell apoptosis*. J Biol Chem, 2004. 279 (46): p. 48197-204.
- 2 - Cao, Y., et al., *Virologic and immunologic characterization of long -term survivors of human immunodeficiency virus type 1 infection*. N Engl J Med, 1995. 332 (4)
- 3 - Campbell G.R et al., *The glutamine-rich region of the HIV-1 tat protein is involved in T-cell apoptosis*. J. Biol Chem, 2004. 279 (46): p. 48197-204.
- 4 - Senkaali D. et al., *The relationship between HIV type 1 disease progression and V3 serotyper in a rural Ugandan cohort*. AIDS reserch and human retroviruses 2004, Vol 20 (9), p.932-937
- 5 - Sharp PM. et al., *The evolution of HIV-1 and the origin of AIDS*. Phil.Trans. R. Soc. B (2010) 365, 2487-2494.
- 6 - UNAIDS/WHO. AIDS: p. 201-8 epidemic update: 2007. Geneva: UNAIDS 2008
- 7 - UNAIDS/WHO. Report on the global AIDS epidemic 2008. Geneva: UNAIDS 2008
- 8 - Hogg R, et al., *Antiretroviral therapy collaboration. Life expectancy of individuals on combination antiretroviral treatment in high income countries: a collaborative analysis of 14 cohort studies*. Lancet 2008; 372: 293-9.
- 9 - Korber B, et al., *Human Retroviruses and AIDS 1998: A compilation and Analysis of Nucleic Acid and Amino Acid Sequences*. Los Alamos National Laboratory Los Alamos, New Mexico.
- 10 - De Marco et al., *Structural Analysis of HIV-1 Maturation Using Cryo-Electron Tomograph*. (2010); *PloS Pathog* 6(11): e1001215. doi:10.1371/journal.ppat.1001215.
- 11 - Giovanni C.D., *Progettazione di inibitori dell'integrasi dell'HIV dotati di elevata potenza e selettività da impiegare nel trattamento dell'AIDS*. (tesi di dottorato,XXI ciclo.2006)
- 12 - Barre-Sinoussi F. et al., *Isolation of a T- lymphotropic retrovirus from a patient at risk for acquired immunodeficiency syndrome (AIDS)*. Science 220: 868-871, 1983.

- 13 - Shaw G.M. et al., *Etiology of AIDS. Virology, molecular biology and evolution of human immunodeficiency viruses*. In: AIDS 1988.
- 14 - Kaplan et al., *Human immunodeficiency virus type 1 virions composed of unprocessed Gag and Gag-Pol precursors are capable of reverse transcribing viral genomic RNA*. Antimicrob Agents Chemother. 1994. 38(12):2929-33.
- 15 - Welman M. et al., *Role of envelope processing and gp41 membrane spanning domain in the formation of human immunodeficiency virus type 1 (HIV-1) fusion-competent envelope glycoprotein complex*. Virus Research Vol.124. 2007. p.103-112.
- 16 - Hauber J. et al., *Trans-activation of human immunodeficiency virus gene expression is mediated by nuclear events*. 1987 . Proc. Natl. Acad.Sci.USA. 84; 6364-6368).
- 17 - Lewer A. et al., *Regulatory protein of HIV. Medical Virology 1*: 155-163, 1991
- 18 - La Placa, cap 67 i retrovirus pag 669.
- 19 - Nie, Z. et al., *The putative alpha helix 2 of human immunodeficiency virus type 1 Vpr contains a determinant which is responsible for the nuclear translocation of proviral DNA in growth-arrested cells*. J Virol, 1998. 72(5): p. 4104-15.
- 20 - Camaur D. et al., *Characterization of human immunodeficiency virus type 1 Vif particle incorporation*. J Virol. 70(9):6106-6111, 1996.
- 21 - Cohen EA, et al., *Role of auxiliary proteins in retroviral morphogenesis*. Curr. Top. Microbiol. Immunol. 214: 219-235, 1996.
- 22 - La Placa, cap. 67, i retrovirus pag 669
- 23 - Clapham and McKnight., *Cell surface receptors, virus entry and tropism of primate lentiviruses*. J Gen Virol. 2002. 83(Pt8): 1809-29.
- 24 - Freed., *HIV-1 replication*. Somat Cell Mol Genet. 2001 26(1-6):13-33.
- 25 - Dimitrov AS., et al., *Role of the fusion peptide and membrane-proximal domain in HIV-1 envelope glycoprotein-mediated membrane fusion*. Biochemistry 2003 42(48):14150-8
- 26 - Pantaleo et al., *Kinetics of human immunodeficiency virus type 1 (HIV-1) DNA and RNA synthesis during primary HIV-1 infection*. Proc Natl Acad Sci U.S.A. 1993. 90(14):6402-9.
- 27 - Kaplan et al ., *The activity of the protease of human immunodeficiency virus type 1*

1994.

28 - Jaskolski M. et al., *Piecing together the structure of retroviral Integrase, an important target in AIDS therapy*. The FEBS Journal, 2009, 276p. 2926-2956.

29 - Ceccherini-Silberstein F. et al., *Characterization and structural analysis of HIV-1 Integrase conservation*. AIDS rev., 2009, 11p.17-29.

30 - Frauke Christ et al., Review *The LEDGF/p75 integrase interaction, a novel target for anti-HIV therapy* Laboratory for Molecular Virology and Gene Therapy, Division of Molecular Medicine, KU Leuven, Kapucijnenvoer 33, 3000 Leuven, Belgium- Virology 435 (2013) 102–109, Elsevier.

31 - Engelman A et al., *Identification of conserved amino acid residues critical of Human Immunodeficiency Virus type 1 Integrase function in vitro*. J. Of Virology, 1992 Vol.66, No.11, p.6361-6369

32 - Esposito D. Craigie, R. *HIV integrase structure and function*. Adv Virus Res 1999, 52, 319-333.

33 - Chiu, T. K.; Davies, D. R. *Structure and function of HIV-1 integrase*. Curr Top Med Chem 2004, 4, 671-686).

34 - Busschots K et al., *Identification of the LEDGF/p75 binding site in HIV-1 integrase*. Journal of Molecular Biology 2007, Vol 354, p.1480-92.

35 - Lodi et al., *Solution Structure of DNA-Binding Domain of HIV-1 Integrase* Biochemistry 1995, Vol.34 p. 9826-9833.

36 - Lutzke, R.A. et al., *Characterization of the minimal DNA-binding domain of HIV integrase protein*. Nucleic Acids Res, 1994. 22(20): p. 4125-31.

37 - Lutzke, R. A. Plasterk, R. H. *Structure-based mutational analysis of the C-terminal DNA-binding domain of human immunodeficiency virus type 1 integrase: critical residues for protein oligomerization and DNA binding*. J Virol 1998, 72, 4841–4848.

38 - Md. Alamgir Hossain et al., *Structural and Functional Insights into Foamy Viral Integrase*. Viruses. 2013, 5(7):1850-1866.

39 - Chiu T.K. Davies D.R. *Structure and function of HIV-1 integrase*. Curr. Top. Med. Chem. 2004, 4, 965-977.

40 - Chen J.C. Et al., *Crystal structure of the HIV-1 integrase catalytic core and C-*

terminal domains: a model for viral DNA binding. Proc. Natl. Acad. Sci. USA 2000, 97, 8233-8238.

41 - Wang J.Y., et al., *Structure a two domain fragment of HIV-1 integrase: implication for domain in the intact protein.* EMBO J. 2001, 20, 7333-7343.

42 - Parril, A.L. *HIV-1 integrase inhibition: binding site, structure activity relationships and future perpesctives.* Curr. Med. Chem. 2003, 10, 1811-1824.

43 - Faure A et al., *HIV-1 Integrase crosslinked oligomers are active in vitro.* Nucleic Acids Research 2005, Vol. 33, N.3, p. 977-986.

44 - Lavanya Krishnan and Alan Engelman. *Retroviral Integrase Proteins and HIV-1 DNA Integration.* JBC Papers in Press, October 5, 2012, DOI 10.1074/jbc.R112.397760.

45 - Engleman A et al., *HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer.* Cell 1991, Vol 67, p. 1211-21.

46 - Chow SA. Et al., *Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus.* Science. 1992;255:723–726.

47 - Delelis O. et al., *Efficient and specific internal cleavage of a retroviral palindromic DNA sequence by tetrameric HIV-1 integrase.* PlosOne, 2007.

48 - Masuda T. et al., *Specific and Independent Recognition of U3 and U5 att Sites by Human Immunodeficiency Virus Type 1 Integrase In Vivo.* Journal of Virology, 1998, V. 72, 10, p. 8396-8402.

49 - Kessl J et al., *HIV-1 Integrase-DNA Recognition Mechanisms.* Viruses 2009, Vol. 1, N.3, p.713-736.

50 - Chen A. et al., *Identification of amino acids in HIV-1 and Avian Sarcoma Virus Integrase subsites required for specific recognition of the Long Terminal Repeat ends.* The Journal of Biological Chemistry, 2006, Vol. 281, No. 7, p. 4173-4182.

51 - Jordan A., et al., *The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation.* Embo J.2001;20:1726–1738.

52 - Delelis O. et al., *Efficient and specific internal cleavage of a retroviral palindromic DNA sequence by tetrameric HIV-1 integrase.* PlosOne, 2007.

53 - Llano M., et al., *Virological and Cellular Roles of the Transcriptional Coactivator*

- LEDGF/p75*. Current Topics in Microbiology and Immunology 2009, Vol. 339, p.125-146.
- 54 - Cereseto A., et al., *Acetylation of HIV-1 integrase by p300 regulates viral integration*. The EMBO Journal 2005, Vol. 7, N.24(17), p.3070-3081.
- 55 - Lin CW. Engleman A. *The barrier-to-autointegration factor is a component of functional human immunodeficiency virus type 1 preintegration complexes*. Virology 2003, Vol.77, p.5030-5036.
- 56 - Ariumi Y et al., *The integrase interactor 1 (INI1) proteins facilitate Tat-mediated human immunodeficiency virus type 1 transcription*. Retrovirology 2006, Vol.3, N.47.
- 57 - Delelis O. et al., *Integrase and integration: biochemical activities of HIV-1 integrase*. Retrovirology, 2008, 5. 114.
- 58 - Barbaro G., et al., *Highly active antiretroviral therapy: current state of the art, new agents and their pharmacological interactions useful for improving therapeutic outcome*. Curr Pharm Des 2005, 11, 1805- 1843.
- 59 - Meadows C. Gervay-Hague J., *Current developments in HIV chemotherapy*. ChemMedChem 2006, 1, 16-29.
- 60 - Frederic Paques and Philippe Duchateau., *Meganucleases and DNA Double-Strand Break-induced recombination: perspectives for Gene Therapy*. Current Gene Therapy, 2007, Vol.7, p. 49-66.
- 61 - Ramirez C.L., et al., *Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects*. Nucleic Acids Reserch 2012, Vol.4, N.12, p.5560-5568.
- 62 - Urnov D. et al., *Genome editing with engineered zinc finger nucleases*. Nature, 2010, Vol. 11.
- 63 - N. Manjunath et al., *Newer Gene Editing Technologies toward HIV Gene Therapy*. Viruses. 2013, 5(11): 2748-2766.
- 64 - Li et al., *In vivo genome editing restores hemostasis in a mouse model of hemophilia*. Nature. 2011, 475(7355):217-221.
- 65 - Handel E.M., et al., *Zinc-Finger nuclease based genome surgery: it's all about specificity*. Current Gene Therapy 2011, Vol.11, N.1, p.28-37.

- 66 - Li L. Chandrasegaran S., *Alteration of the cleavage distance of Fok I restriction endonuclease by insertion mutagenesis*. Proc Natl Acad Sci U S A. 1993, 90(7):2764-8.
- 67 - Jurate Bitinaite., et al., *FokI dimerization is required for DNA cleavage*. June 22, 1998.
- 68 - David A. Wah et al., *Structure of the multimodular endonuclease FokI bound to DNA*. Nature 1997, 388(6637): 97-100.

**ANALISI *IN SILICO* DI SEGNALI DI LOCALIZZAZIONE
NUCLEARE SUL PROTEOMA DI CHLAMYDIA SPP. E STUDIO
FUNZIONALE DI DUE NLS DELLA PROTEINA EFFETTRICE
PMPD**

FAMIGLIA CHLAMYDIACEAE

1. INTRODUZIONE

La famiglia Chlamydiaceae è composta da batteri gram-negativi intracellulari obbligati capaci di infettare una vasta gamma di cellule eucariotiche. ^[1]

A questa famiglia di batteri appartengono le specie: *C. pneumoniae*, *C. pecorum*, *C. felis*, *C. psittaci*, *C. abortus*, *C. caviae*, *C. muridarum*, *C. trachomatis* e *C. suis*, tutti con un ciclo di sviluppo simile ma molto diversi dal punto di vista del tropismo tissutale. Infatti *C. trachomatis* infetta l'epitelio genitale e la mucosa oculare dell'uomo, *C. pneumoniae* infetta il tratto respiratorio umano, *C. psittaci* infetta prevalentemente gli uccelli, *C. caviae* infetta la congiuntiva dei porcellini d'india e *C. muridarum* è un patogeno murino.

Le Clamidiose hanno un'organizzazione cellulare simile a quella dei batteri Gram-negativi. La parete cellulare è, infatti, composta da una membrana interna ed una esterna, separate da uno spazio intermembrana detto periplasma. A differenza, però, degli altri batteri Gram-negativi, le Clamidiose possiedono un lipopolisaccaride (LPS) tronco, una serie di porine indicate con l'acronimo MOMP (*Major Outer Membrane Proteins*) e un peptidoglicano (PNG) non sempre rilevabile in quanto normalmente è sostituito da uno strato di proteine ricche in cisteina (CRP o *Cysteine-Rich Proteins*).

Il ciclo vitale delle Clamidiose è dimorfico (figura 1), caratterizzato dall'alternanza di due forme morfologiche distinte: un corpo elementare (CE) e un corpo reticolare (CR).

Il CE è di microscopiche dimensioni (0,2 µm), rotondo, con il citoplasma denso e compatto e rappresenta la forma infettante. Esso è privo di capacità replicativa ma resistente e stabile anche al di fuori dell'ambito cellulare, infatti è in questa forma che avviene l'interazione tra proteine batteriche di membrana (GAG, MOMP, OmcB e

PmpD) e recettori della cellula ospite (glucani, eparani). [2] Tali interazioni permettono l'endocitosi delle Clamidio da parte della cellula parassitata.

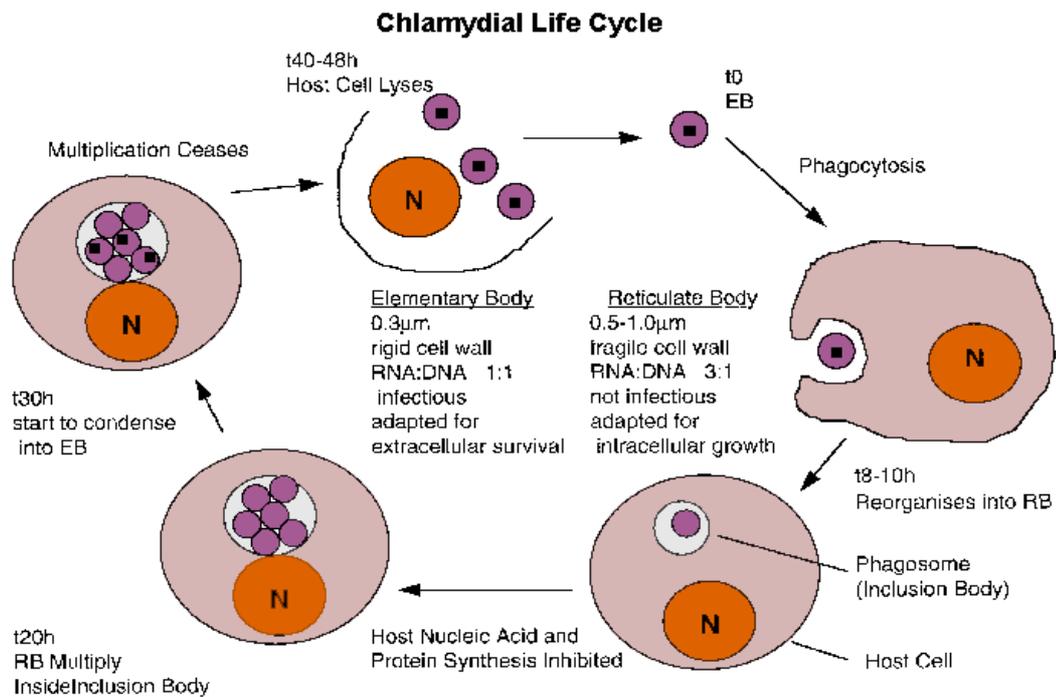


Figura1: Ciclo vitale di clamidia

All'interno della cellula ospite il CE va incontro ad un processo di riorganizzazione che si manifesta con l'aumento di volume e la trasformazione in corpi reticolari (CR).

Il CR è di dimensioni maggiori (0,5-1,0 µm) rispetto al CE, si moltiplica per scissione binaria all'interno delle inclusioni, non è resistente in ambiente extracellulare e non è infettante. [3]

Circa a metà del ciclo d'infezione i CR si reversiono in CE che vengono rilasciati all'esterno della cellula ospite per lisi cellulare o fusione dell'inclusione con la membrana cellulare. [4]

Per quanto riguarda il genoma delle clamidio, esso è costituito da un cromosoma

circolare di circa 1,0-1,1 Mbp e i geni, durante l'infezione, vengono espressi a cascata: dapprima i geni definiti early (1-8 h), seguiti dai geni mid (8-16 h) ed infine i geni late (16-24).^[5]

Queste ondate di produzione sembrano coordinate con il ciclo d'infezione in modo da modulare le funzioni cellulari necessarie allo sviluppo delle clamidie. A questo scopo, le proteine derivanti dall'ondata tardiva, “impacchettate” nel CE al momento della sua formazione, vengono liberate nel citoplasma della cellula, non appena il batterio vi si attacca. Questi effettori determinano l'ingresso del parassita, disarmano la risposta immunitaria naturale e ritardano la maturazione dell'endosoma. La seconda ondata di proteine effettrici, comprese alcune Incs, partecipano alla biogenesi dell'inclusione, determinano l'acquisizione di nutrienti, manipolano il ciclo cellulare ed innescano il processo dell'infiammazione.^[6] Infine, la terza ondata di proteine prepara il corredo proteico contenuto nei corpi elementari e determina l'uscita dell'inclusione dall'ospite.

L'analisi comparativa dei genomi delle varie specie di *Chlamydia* ha permesso di identificare due gruppi di proteine non riscontrabili in altri microrganismi: le “*Inclusion Membrane Proteins*” (Incs) e le “*Polymorphic Membrane Proteins*” (Pmps).^[7]

Le Incs presentano un ampio motivo idrofobico e bilobato, di circa 50-80 aminoacidi.^[8] Sono proteine inserite nella membrana dell'inclusione grazie al sistema di secrezione di tipo III, tipico dei batteri Gram- negativi^[9] e non hanno un peptide segnale. Alcune proteine Inc partecipano alla biogenesi dell'inclusione e alla modulazione delle funzioni cellulari, ma per la maggior parte di esse non si è ancora definita una funzione.

Le Pmps, proteine ad elevato peso molecolare, sono esposte sulla superficie esterna del batterio.

la loro funzione, sembra, sia alla base di un meccanismo di variazione antigenica, che permette alle clamidie di evadere le difese dell'ospite.^[10, 11] Tutte le proteine appartenenti a questa famiglia condividono ripetizioni multiple dei motivi GGAI (L o V) e FXXN (dove la X rappresenta qualsiasi aminoacido), localizzate per lo più nella regione N-terminale della proteina.^[12, 13]

le Pmps, sono proteine autotrasportatrici costituite da tre domini: una *sequenza segnale* (SS) all'N-terminale della proteina, un “*passenger domain*” (PD) con funzione effettrice, e un *dominio di traslocazione* al C-terminale (TD), che forma un poro con struttura a *beta-barrel* la cui funzione è quella di facilitare la traslocazione del PD sulla superficie del batterio. ^[14]

I geni pmps sono organizzati in due clusters, uno contenente il gruppo dei geni A, B e C, e l'altro contenente i geni E, F, G, H ed I. Il gene D, invece, risulta essere isolato ed è in singola copia insieme al gene A ed H, tutti gli altri sono presenti in più copie.

Ultimamente si è venuti a conoscenza di proteine effettrici del genere Chlamydia in grado di attraversare la membrana nucleare e modulare le funzioni cellulari agendo direttamente sul materiale genico della cellula ospite. Infatti Pennini et al, ^[15] hanno identificato una proteina effettrice, detta NUE, che nella fase tardiva dell'infezione, viene secreta dal microrganismo e raggiunge il nucleo della cellula ospite, dove è in grado di legare la cromatina e metilarne gli istoni. Questa traslocazione è facilitata da due segnali di localizzazione nucleare (NLS) all'interno della proteina che, se deleti, diminuiscono fortemente l'importo nucleare della stessa. Alla luce di ciò, abbiamo effettuato un'analisi *in silico* nel genere clamidia per identificare altre proteine presentanti segnali di localizzazione nucleare e studiarne la funzione. Dall'analisi è risultato che ben 72 proteine presentano uno o due NLS putativi e tra queste proteine rientra anche la PmpD, nota come autotrasportatore e già studiata da Swanson et al. per essere una delle proteine più immunogene di Chlamydia e pertanto interessante per lo sviluppo di un vaccino. ^[14] Gli NLS identificati su questa proteina sono:

- un NLS bipartito classico (aa 783-798), costituito da due cluster di residui basici separati da 9 aa (KRR₉KRVR);
- un NLS putativo non canonico (aa 1377-1392), identificato da noi mediante l'analisi *in silico*. Questo NLS è costituito da un motivo PY associato ad un cluster di residui basici verso l'N-terminale (VGPV_{RPTFYAL}H_{FN}PY).

Mio contributo a questo lavoro è stato quello di caratterizzare la funzionalità dell'NLS canonico legato al frammento proteolitico p30 segnalato da Swanson et al.

2. MATERIALI E METODI

COSTRUTTI UTILIZZATI NEL SEGUENTE LAVORO:

- **pEPI-GFP + pmpD p30 (762-1071).** L'amplificazione del gene p30 NLS pmpD (762-1071) è stata ottenuta utilizzando come template il genoma del ceppo GO/86 di *C. trachomatis*, e come primer le sequenze nucleotidiche:

p30 NLS pmpD Up CACCCTTTTCGCATCTGAAGATGGG (sequenza d'orientamento)
p30 NLS pmpD Lo TTATTGAGCATTCTTCGCAATCCAA (STOP codon)

Il prodotto della reazione (30 cicli con temperatura di annealing di 56°C) è stato separato elettroforeticamente su gel di agarosio, estratto, purificato, quindi trasferito nel vettore pET160/GW/D-TOPO (Invitrogen, Carlsbad, CA, USA).

Successivamente, la proteina d'interesse è stata trasferita nel vettore pEPI-GFP mediante tecnologia *Gateway*, per generare il costrutto pEPI-GFP + pmpD p30 (762-1071). Il vettore pEPI-eGFP, derivato dal plasmide commerciale pGFP-C1 (Clontech), è un generoso dono del Prof. David Jans (Monash University, Australia) è un vettore di espressione eucariotica che consente l'espressione di proteine ricombinanti in fusione N-terminale con la green fluorescent protein.

- **pEGFP-C1.** Il vettore pEGFP-C1 è un vettore di espressione eucariotica, codificante la proteina fluorescente GFP, sotto l'azione del promotore *immediate early* di HCMV. Tale vettore è stato utilizzato nei saggi di trasfezione come controllo negativo.
- **pcDNA 3.1 mCherry + UL42 NLSwt:** è un costrutto, da noi utilizzato come controllo positivo nei saggi di trasfezione, che esprime la proteina mCherry in fusione con l'NLS bipartito di UL42, il fattore di processività della DNA polimerasi di HSV-1. ^[16]

- **pcDNA 3.1 mCherry + pmpD NLSwt (783-798):** è un costrutto che esprime la proteina mCherry in fusione con l'NLS bipartito del prodotto proteolitico p30 della proteina pmpD di *C. trachomatis*. Anche tale costrutto è stato utilizzato come controllo positivo nei saggi di trasfezione.

CELLULE.

Le cellule A549 (cellule neoplastiche derivanti da un carcinoma bronchiolo-alveolare di un uomo caucasico di 58 anni) e FM-hMSC (*Fetal Membrane-human Mesenchymal Stem Cells*: cellule staminali mesenchimali isolate da placente a termine di madri donatrici sane) sono state utilizzate nel presente lavoro. Le cellule sono state mantenute in coltura in *Dulbecco's Modified Eagle's Medium* (DMEM) supplementato con il siero fetale bovino, 5% (v / v) per le A549 e 10% (v / v) per le FM-hMSC, 50 unità / ml di penicillina, 50 unità / ml streptomicina, e 2 mM L-glutammina.

Per eseguire gli esperimenti di localizzazione subcellulare, le cellule sono state tripsinizzate e seminate su vetrini in piastre da 24 pozzetti 1 giorno prima della trasfezione. La trasfezione è stata eseguita utilizzando il kit *TransIT-LTI* (Mirus, Madison, WI), secondo le indicazioni della casa produttrice.

ANALISI DELLA LOCALIZZAZIONE DEL SEGNALE.

Per l'acquisizione delle immagini, i vetrini con le cellule trasfettate sono stati fissati in Paraformaldeide 4%, dopo 48 e 72 h dalla trasfezione, e fotografati utilizzando un microscopio a fluorescenza con ottica invertita Nikon Eclipse TE2000U (Nikon, Tokio, Japan).

Le immagini digitali sono stati ottenute utilizzando il software NisElement AR 2.30 (Nikon Corp., Tokyo, Giappone).

A fini statistici sono state analizzate le immagini di almeno 30 cellule trasfettate. Di ogni

cellula, è stata valutata la fluorescenza del nucleo, del citoplasma e del rumore di fondo mediante l'ausilio del software ImageJ (National Institute of Health, USA). Per l'attendibilità dei dati ciascun esperimento è stato ripetuto tre volte.

Dai valori ottenuti è stato estrapolato un coefficiente di importo al nucleo, dato dal rapporto tra la fluorescenza nucleare (FN) e quella citoplasmatica (FC), sottraendo ad entrambe il rumore di fondo.

In base al valore ottenuto è stato possibile stabilire se il segnale è ubiquitario, nucleare o più nucleare che citoplasmatico. Nel primo caso il valore del coefficiente medio di importo al nucleo (N/C) è al di sotto di 2; nel caso di un segnale fortemente nucleare è stato ottenuto un valore (N/C) superiore a 4, e se il segnale presenta una localizzazione più nucleare che citoplasmatica il valore (N/C) è compreso fra 2 e 4.

3. RISULTATI

Allo scopo di verificare la localizzazione intracellulare della proteina pmpD p30 (762-1071), le cellule A549 e FM-hMSC sono state trasfettate con il costrutto pEPI-GFP-N1 + pmpD p30 (in fusione N-terminale con GFP) e in contemporanea sono stati utilizzati in ciascun saggio, come controllo negativo, il costrutto pEGFP-C1 e come controlli positivi il costrutto pcDNA 3.1 mCherry + UL42 NLSwt, contenente un NLS di HSV-1 precedentemente caratterizzato in laboratorio, ^[16] e pcDNA 3.1 mCherry + pmpD NLSwt, contenente l'NLS canonico presente sul frammento p30.

Alle 72 ore *post transfectionem* le cellule sono state fissate e la localizzazione delle varie proteine espresse è stata evidenziata mediante un microscopio a fluorescenza con ottica invertita.

Come previsto, il segnale del prodotto di pEGFP-C1 (controllo negativo) è stato rilevato sia nel nucleo che nel citoplasma, infatti il valore del coefficiente medio d'importo al nucleo (N/C) è pari a 1,6 nelle cellule A549 e 1,5 nella linea cellulare FM-hMSC.

Per quanto riguarda i risultati dei controlli positivi: la localizzazione intracellulare del prodotto di pcDNA 3.1 mCherry + UL42 NLSwt è stata interamente ristretta al nucleo, presenta infatti un N/C pari a 11,88 in A549 e 7,67 in FM-hMSC, confermando quindi i risultati precedentemente ottenuti in laboratorio. Anche il prodotto di pcDNA 3.1 mCherry + pmpD NLSwt (783-798) è stato rilevato più nucleare che citoplasmatico, difatti il valore N/C è pari a 3,25 in A549 e 3,40 in FM-hMSC.

L'analisi della proteina ricombinante codificata da pEPI-GFP-N1 + pmpD p30 (762-1071) ha mostrato una distribuzione del segnale ubiquitaria all'interno delle cellule analizzate, nello specifico in A549 il valore N/C è risultato essere pari a 1,2. mentre in FM-hMSC il valore N/C calcolato è stato di 1,6.

I dati ottenuti sono stati utilizzati per generare grafici mediante il software *Excel* (Microsoft, USA).

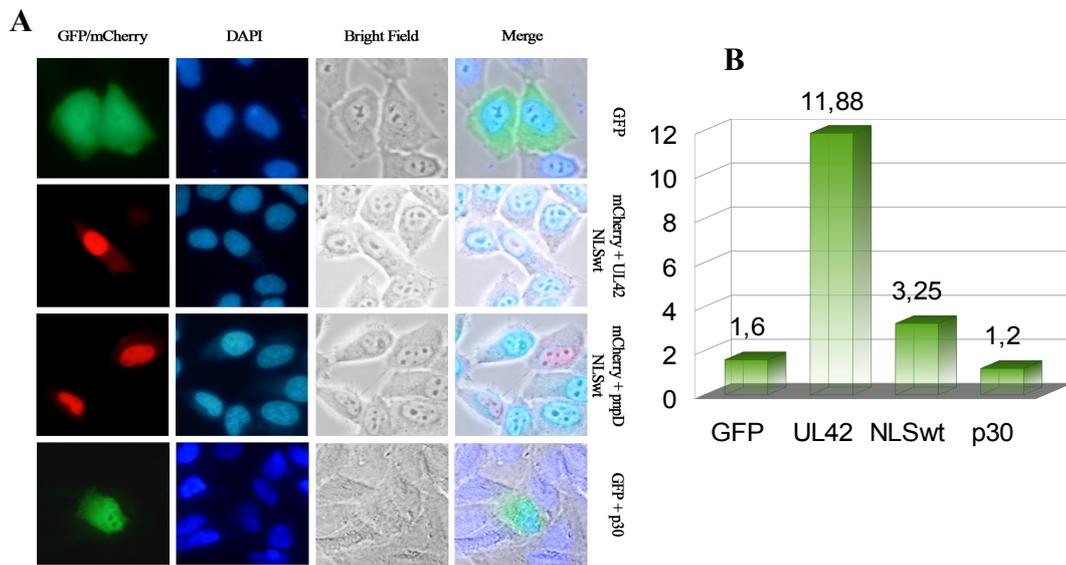


Figura 2. A: analisi, mediante microscopia ad EPI-fluorescenza, della localizzazione sub-cellulare di GFP, UL42, NLSwt e p30 in cellule A549. **B:** rappresentazione grafica dei valori medi dei coefficienti di importo al nucleo ottenuti.

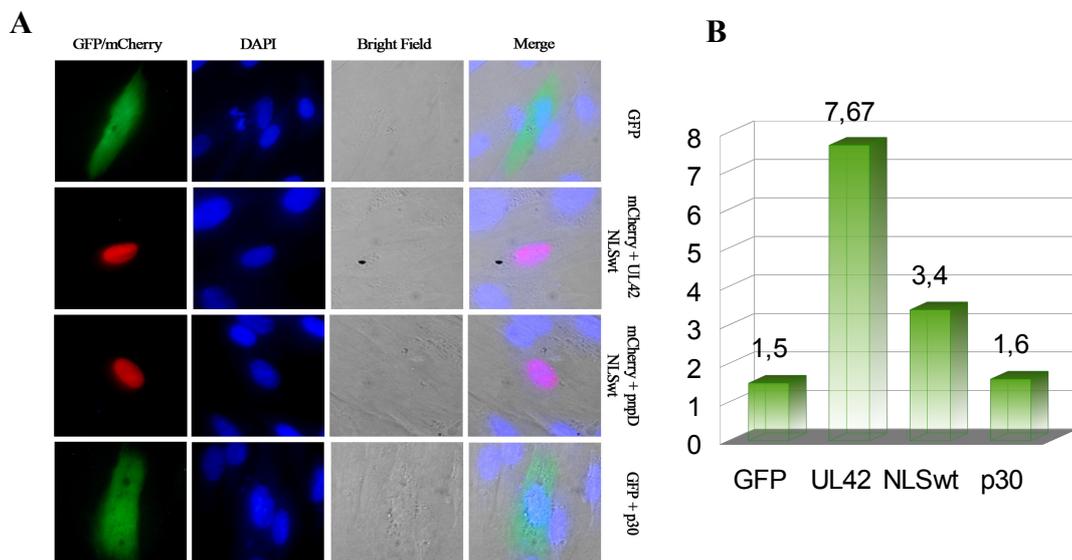


Figura 3. A: analisi, mediante microscopia ad EPI-fluorescenza, della localizzazione sub-cellulare di GFP, UL42, NLSwt e p30 in cellule FM-hMSC. **B:** rappresentazione grafica dei valori medi dei coefficienti di importo al nucleo ottenuti.

4. CONCLUSIONI

I batteri patogeni utilizzano una vasta gamma di strategie atte a evitare l'eliminazione dal loro ospite. E' noto infatti che le clamidie, attraverso meccanismi ancora poco chiari, inibiscono l'apoptosi nelle fasi precoci dell'infezione, e promuovono la morte programmata cellulare in fase tardiva, per poter fuoriuscire e diffondere nell'ambiente esterno. ^[17] E' stato ampiamente dimostrato che molte proteine effettrici della clamidia modulano diverse attività all'interno del citoplasma delle cellule ospiti, mentre per quanto riguarda l'attività all'interno del nucleo solo di recente è stata identificata da Pennini et al., una proteina effettrice, detta NUE, che nella fase tardiva dell'infezione, viene secreta dal microrganismo e raggiunge il nucleo della cellula ospite, dove è in grado di legare la cromatina e metilarne gli istoni. Questa traslocazione è facilitata da due segnali di localizzazione nucleare (NLS) all'interno della proteina che, se deleti, diminuiscono fortemente l'importo nucleare della stessa.

Dal momento che le proteine codificate da batteri che localizzano nei nuclei delle cellule rappresentano un campo innovativo per capire meglio l'interazione ospite-patogeni, abbiamo effettuato un'analisi *in silico* nel genere clamidia per identificare altre proteine presentanti segnali di localizzazione nucleare e studiarne la funzione. Dall'analisi, effettuata da colleghi, è risultato che ben 72 proteine presentano uno o due NLS putativi e tra queste proteine rientra anche la PmpD, già studiata da Swanson et al. e definita come autotrasportatore. ^[14] Questa proteina presenta due NLS: il primo è un NLS classico bipartito, inserito nel “*passenger domain*” (PD), l'effettore secreto dalla proteina pmpD; il secondo NLS, identificato dai colleghi, è un NLS PY che è incorporato nel dominio di traslocazione di pmpD. E' stato dimostrato anche che i due NLS sono in grado di traslocare una proteina cargo nel nucleo con coefficienti di importo nucleare rispettivamente di 3,25 e 4,64.

Mio contributo a questo progetto è stato quello di valutare se l'NLS classico è in grado di traslocare il frammento p30 all'interno del nucleo ed eventualmente capirne gli effetti. A

tal fine, abbiamo verificato la localizzazione intracellulare del complesso p30-NLS, mediante la sua espressione in fusione con il marcatore fluorescente GFP, in cellule A549 e FM-hMSC.

Come controllo di localizzazione ubiquitaria è stata utilizzata la proteina GFP, mentre come controllo di localizzazione nucleare è stata utilizzata sia la proteina virale UL42wt che l'NLS di pmpD, entrambe in fusione con mcherry.

I risultati ottenuti dai saggi indicano che l'NLS classico legato alla porzione p30, si diffonde ubiquitariamente all'interno delle cellule A549 (N/C= 1,2) e FM-hMSC (N/C= 1,6).

E' stato dimostrato che la pmpD, una volta espressa, viene esposta sulla membrana e può andare incontro a un processo proteolitico. Tale processo è strutturato in due fasi: una precoce, in cui vengono generati frammenti associati alla membrana (p155, p73 e p82), ed una tardiva, ad almeno 30-36 ore dall'infezione, in cui si evidenziano frammenti solubili (p111, p73 e p30).^[18] Quindi dal momento che la p30 viene generata nella fase tardiva, il nostro risultato potrebbe indicare una sua funzionalità a livello citoplasmatico mentre è l'altra porzione di pmpD, la p82, ad essere traslocata a livello nucleare della cellula infetta. La p82 presenta ambedue gli NLS ed inoltre viene prodotta nella fase precoce. Per appurare questa ipotesi in questi giorni stiamo analizzando la localizzazione intracellulare della proteina p82.

Questo studio ha come intento principale quello di puntare l'attenzione verso un insieme di prodotti proteici espressi dalla clamidia che, in virtù del loro essere corredati di NLS, potrebbero raggiungere il nucleo della cellula ospite. In tale sede potrebbero interferire con l'espressione di geni cellulari con varie di modalità, non ultima quella epigenetica. Le conseguenze di attività di questo genere si potrebbero ripercuotere sulla stabilità genetica delle cellule infettate in quei tessuti in cui le clamidie tendono a stabilire infezioni croniche, e alla lunga contribuire allo sviluppo di tumori secondo meccanismi finora non sospetti.

5. BIBLIOGRAFIA

- 1 – Kalayoglu MV., *Ocular chlamydial infections: pathogenesis and emerging treatment strategies*. Curr Drug Targets Infect Disord. 2002; 2(1):85-91.
- 2 – Dautry-Varsat., et al., *Recent insights into the mechanisms of Chlamydia entry*. Cell Microbiol, 2005. 7(12): p. 1714-22.
- 3 - Moulder J.W., et al., *Interaction of chlamydiae and host cells in vitro*. Microbiol Rev, 1991. 55(1): p. 143-90.
- 4 - Brunham R. C. and Rey-Ladino J., *Immunology of Chlamydia infection: implications for a Chlamydia trachomatis vaccine*, Nature Reviews Immunology, Feb. 2005: p. 149-161.
- 5 - Belland R.J., et al., *Genomic transcriptional profiling of the developmental cycle of Chlamydia trachomatis*. Proc Natl Acad Sci U S A, 2003. 100(14): p. 8478-83.
- 6 - Buchholz K.R. and R.S. Stephens., *The extracellular signal-regulated kinase/mitogen-activated protein kinase pathway induces the inflammatory factor interleukin-8 following Chlamydia trachomatis infection*. Infect Immun, 2007. 75(12): p. 5924-9.
- 7 - Rockey D. D. et al., *Genome Sequencing and Our Understanding of Chlamydiae*, Infection and Immunity, Oct. 2000: p. 5473-5479.
- 8 - Bannantine J.P., et al., *A secondary structure motif predictive of protein localization to the chlamydial inclusion membrane*. Cell Microbiol, 2000. 2(1): p. 35-47.
- 9 - Su H., et al., *A recombinant Chlamydia trachomatis major outer membrane protein binds to heparan sulfate receptors on epithelial cells*. Proc Natl Acad Sci U S A, 1996. 93(20): p. 11143-8.
- 10 - Rockey D. D. et al., *Genome Sequencing and Our Understanding of Chlamydiae*, Infection and Immunity, Oct. 2000: p. 5473-5479.
- 11 - Bavoil P. and P. Wyrick., *Chlamydia: genomics and pathogenesis*, 2006, Oxford: Horizon Bioscience. XII: p. 542.
- 12 - Rockey, D.D., et al., *Cloning and characterization of a Chlamydia psittaci gene*

coding for a protein localized in the inclusion membrane of infected cells. Mol Microbiol, 1995. 15(4): p. 617-26.

13 - Sriprakash K.S and E.S. MacAvoy., *A gene for dnaB like protein in chlamydial plasmid. Nucleic Acids Res*, 1987. 15(24): p. 10596.

14 - Swanson K. A. et al., *Chlamydia trachomatis polymorphic Membrane Protein D is an Oligomeric Autotransporter with a Higher-Order Structure, Infection and Immunity*, Jan. 2009. 77(1): p. 508-516.

15 - Pennini ME et al., *Histone Methylation by NUE, a Novel Effector of the Intracellular Pathogen Chlamydia trachomatis. Plos Pathog*, 2010. 6(7): e1000995.

16 - Alvisi G. et al., *Nuclear Import of HSV-1 DNA Polymerase Processivity Factor UL42 Is Mediated by a C-Terminally Located Bipartite Nuclear Localization Signal, Biochemistry*, 2008. 47(52): p. 13764-13777.

17 - Byrne G.I. and D.M. Ojcius, *Chlamydia and apoptosis: life and death decisions of an intracellular pathogen. Nat Rev Microbiol*, 2004. 2(10): p. 802-8.

18 - Sun S.C. and S.C. Ley., *New insights into NF-kappaB regulation and function. Trends Immunol*, 2008. 29(10): p. 469-78.