

Università degli Studi di Napoli "Federico II"
Dipartimento di Chimica Farmaceutica e
Tossicologica
Facoltà di Farmacia



Dottorato di Ricerca

XXI ciclo

**PROGETTAZIONE DI INIBITORI DELL'INTEGRASI
DELL'HIV-1 DOTATI DI ELEVATA POTENZA E
SELETTIVITÀ DA IMPIEGARE
NEL TRATTAMENTO DELL'AIDS**

Tutors
Ch.mi Proff.
Ettore Novellino
Antonio Lavecchia

Candidata
Dott.ssa
Carmen Di Giovanni

Coordinatore
Ch.ma Prof.ssa
M.Valeria D'Auria

Indice

Indice	1
Parte I	4
Introduzione	4
Cenni storici	6
La XIII Conferenza internazionale mondiale.....	7
Il “Medical Act”	9
L’AIDS nel mondo	11
Il virus dell’HIV.....	13
Virus HIV-1 e HIV-2	14
Struttura del virione.....	15
Il genoma di HIV	20
Proteine strutturali di HIV	25
Poliproteina Pol	28
Poliproteina Env	33
Poliproteina Gag.....	36
Proteine regolatrici ed accessorie di HIV.....	39

Infezione e ciclo vitale di HIV.....	46
Infezione da HIV dei linfociti	53
Patogenesi di HIV.....	55
Infezione acuta o primaria.....	56
Latenza clinica	58
Fase sintomatica o AIDS conclamato	59
La diagnosi di infezione da HIV	62
La trasmissione del virus	69
Sintomatologia dell'AIDS	71
La terapia farmacologica.....	72
Il fallimento terapeutico.....	77
La Farmacoresistenza	79
Parte II	84
Descrizione del target: l'enzima integrasi di HIV-1	84
Modelli di interazione dell'integrasi dell'HIV-1 con il DNA.....	93
Meccanismo d'azione dell'integrasi	96
Cenni di letteratura: progettazione del 5-CITEP e dei primi ligandi attivi contro l'integrasi di HIV-1	100
Mercaptosalicilidrazidi	108
Composti ossidrilati.....	109

Arildichetoacidi.....	119
Parte III.....	154
Progettazione di inibitori dichetoacidi bifunzionali (BDK)	154
Saggi di attività biologica contro HIV-1	164
Saggi di citotossicità.....	165
Studi di molecular modeling.....	168
Metodi computazionali.....	176
Parte IV	181
Progettazione e sintesi di inibitori dichetoacidi monofunzionali (DKA)	
.....	181
Saggi di attività biologica verso HIV-1	184
Saggi di citotossicità.....	187
Molecular modeling.....	188
Metodi computazionali.....	197
Conclusioni	204
Bibliografia.....	207
Siti Internet consultati.....	220

Parte I

Introduzione

La Sindrome da ImmunoDeficienza Acquisita (SIDA) o Acquired Immune Deficiency Syndrome (AIDS), è un insieme di manifestazioni dovute alla deplezione di linfociti T. È caratterizzata dalla progressiva compromissione delle difese immunitarie e dall'insorgenza di gravi patologie, come cancro (frequente è il sarcoma di Kaposi) o encefaliti, oppure dalla comparsa di infezioni opportuniste che si sviluppano nei pazienti debilitati dalla malattia. L'individuo infettato dal virus diventa portatore asintomatico ed è detto sieropositivo perché nel suo sangue è possibile riscontrare la presenza di anticorpi anti-HIV; può sviluppare in seguito la sindrome vera e propria (AIDS conclamata). Quando questa compare, provoca un rapido deperimento fisico; l'esito dell'AIDS conclamata è infausto. Il termine AIDS è l'acronimo di Acquired ImmunoDeficiency Syndrome, sindrome da immunodeficienza acquisita. In queste manifestazioni sono comprese infezioni da microrganismi rari o non patogeni ed insorgenza di tumori sia comuni nella popolazione generale sia

caratteristici delle persone immunocompromesse sia peculiari di chi presenta tale sindrome. numerosi studi hanno verificato e confermato il legame tra il virus HIV e l'AIDS, accettato dalla comunità scientifica¹. La sindrome è, allo stato attuale, curabile con numerosi farmaci ma ancora non guaribile, nel senso che non è possibile eradicare totalmente il virus dall'ospite. Le terapie odierne, di gran lunga meglio tollerate di quelle usate al momento dell'emergenza dei primi anni '80, riescono a ridurre la viremia (quantità di virus presente nel sangue) a livelli bassissimi o non rilevabili consentendo la rigenerazione dei linfociti e la prosecuzione di una vita esente dalle malattie opportunistiche che normalmente si presentano nelle persone non curate. L'andamento clinico-patologico della sindrome è estremamente variabile tra gli individui per il fatto che la progressione dell'infezione dipende da fattori genetici sia del virus che dell'ospite² che dalle condizioni igieniche e dalle co-infezioni³. Nei paesi in cui le costose cure antiretrovirali e le cure per le infezioni opportunistiche e neoplastiche sono maggiormente disponibili, o come in Italia pagate dal SSN, la mortalità dell'AIDS è di molto ridotta, bilanciata però dai problemi causati dagli effetti collaterali⁴ dallo sviluppo di resistenza ai farmaci, dalla scarsa aderenza ai regimi terapeutici prescritti.

Cenni storici

Gli scienziati hanno identificato un tipo di scimpanzè in Africa occidentale come fonte di infezione dell'HIV negli esseri umani. Il virus molto probabilmente è passato da una specie all'altra quando gli uomini africani hanno utilizzato la carne di questi



animali per cibarsi, entrando in contatto con il loro sangue infetto. In molti anni, il virus si è diffuso lentamente attraverso l'Africa e successivamente in altre parti del mondo.

Nel 1983 il medico e virologo francese Luc Montagnier⁵⁻⁶ ed altri scienziati dell'Institut Pasteur di Parigi isolarono, dal linfonodo di un uomo a rischio di sviluppare l'AIDS, quello che sembrava essere un nuovo retrovirus umano. Poco tempo dopo, sia il gruppo di Robert Gallo al National Cancer Institute di Bethesda, nel Maryland, sia il gruppo guidato da Jay Levy all'Università della California a San Francisco, isolarono un retrovirus da persone infette dal virus ma non malate (denominate sieropositive) e da pazienti con AIDS conclamata (ovvero, che manifestavano già il quadro clinico della sindrome). Tutti e tre i gruppi avevano così isolato quello che oggi è noto come HIV.

La XIII Conferenza internazionale mondiale

Dal 1985 si tiene ogni anno una Conferenza internazionale sull'AIDS nella quale tutti gli aspetti relativi all' infezione da HIV sono affrontati dal punto di vista biologico, medico, farmaceutico, politico e sociale. La XIII Conferenza mondiale si è tenuta nel luglio 2000 a Durban, in Sudafrica. Tra i risultati del meeting, particolarmente rilevanti sono stati quelli relativi alla trasmissione materno-fetale del virus, per la quale un nuovo farmaco, la nevirapina, risulterebbe più efficace della terapia a base di AZT e 3TC. La trasmissione materno-fetale (o verticale) causa ogni anno la nascita di 600.000 bambini sieropositivi.

Durante il convegno, 5228 medici e ricercatori di 84 paesi hanno siglato la Dichiarazione di Durban, in cui si afferma che, in tutto il mondo, i soggetti infetti da HIV (sieropositivi e malati di AIDS) sono 34 milioni, di cui 24 milioni si trovano nell'Africa subsahariana; nel 1999 2,6 milioni di persone sono morte per AIDS, il valore più elevato dall'inizio dell'epidemia; nel documento, inoltre, si ribadisce la natura virale dell'infezione, precedentemente messa in discussione da alcuni scienziati.

Questione centrale del convegno è stata anche la possibilità di accesso ai farmaci brevettati per i paesi in via di sviluppo, nei quali le precarie condizioni igienico-sanitarie e la mancanza di adeguate condizioni economiche precludono la possibilità di comprare i farmaci protetti da brevetto per i trattamenti antivirali e, di fatto, aumentano la percentuale di decessi dovuti ad AIDS. La World Trade Organization (WTO) ha riconosciuto ai paesi in via di sviluppo fino al gennaio del 2000 (ed ai paesi più poveri fino al 2006) la possibilità di utilizzare farmaci brevettati senza il consenso dell'industria che ne detiene il brevetto (principio del compulsory licensing) e la possibilità di importare i farmaci dal paese in cui vengono venduti al prezzo più basso (parallel import). L'utilizzo di farmaci detti "generici", più economici perché non soggetti a brevetto e formulati con gli stessi principi attivi dei prodotti brevettati, ha permesso ad alcuni paesi come il Brasile di ridurre notevolmente i costi delle terapie.

Per sollecitare l'attenzione del grande pubblico sui progressi della ricerca nella lotta all'AIDS, sullo stato dell'epidemia nel mondo e sulle iniziative adottate da varie organizzazioni, ogni anno, il giorno 1° dicembre, ricorre la Giornata mondiale dell'Aids, istituita per la prima volta nel 1988 dall'OMS.

Il “Medical Act”

In Sudafrica, in cui si stimano 5 milioni di sieropositivi, al momento della XIII Conferenza era in corso una disputa giudiziaria che da quattro anni opponeva il governo sudafricano a trentanove multinazionali farmaceutiche. Queste, infatti, avevano citato in giudizio il governo di Pretoria che, nel 1997, nell'intento di assicurare alle centinaia di migliaia di persone colpite dall'AIDS l'accesso alle cure fondamentali, aveva siglato il documento noto come “Medical Act” in cui autorizzava la produzione interna e l'acquisto presso altri paesi di farmaci generici antivirali. Nell'aprile 2001, accogliendo l'appello di Nelson Mandela, le trentanove multinazionali hanno ritirato la denuncia. All'appello di Mandela avevano aderito importanti organizzazioni internazionali (tra cui Médecins sans Frontières ed Amnesty International) ed erano state raccolte in pochi mesi centinaia di migliaia di firme in tutto il mondo

Tra i molteplici appuntamenti internazionali relativi all'AIDS, di particolare interesse è stata la XII Conferenza internazionale sull'Aids e malattie sessualmente trasmissibili in Africa (ICASA), svoltasi dal 9 al 13 dicembre 2001 a Ouagadougou, in Burkina. In questa sede sono emersi nuovi preoccupanti dati sull'espansione dell'infezione: il numero dei sieropositivi nel mondo è stimato in

40 milioni, il 75% dei quali vive in Africa. Di questa quota, il 20-25% si trova in Sudafrica, il 10-15% nell'Africa centrale e nella regione dei Grandi Laghi, mentre percentuali inferiori sono distribuite tra il Nordafrica e l'Africa occidentale. È stata sottolineata la necessità di attuare una profilassi di tipo vaccinale, l'urgenza di mobilitare le risorse economiche a questo scopo e di tenere sotto controllo la trasmissione dell'HIV tra madre e figlio nel corso della gravidanza; inoltre, è stata riportata un'interessante sperimentazione guidata dal medico francese Luc Montagnier in Costa d'Avorio, basata sulla somministrazione per sei mesi di una combinazione di tre farmaci fino a ridurre e stabilizzare la viremia, seguita dall'assunzione di estratti naturali di papaia. Obiettivo della ricerca è di trovare cure alternative che rendano più tollerabili quelle tradizionali e possano eventualmente sostituire queste ultime, almeno in parte.

La XIV Conferenza internazionale sull'AIDS mondiale si è tenuta dal 7 al 12 luglio 2002 a Barcellona. La catastrofica previsione emersa dal rapporto della UNAIDS, in base al quale da oggi al 2020 i morti per AIDS potrebbero ammontare a 68 milioni, è stata in parte bilanciata da nuove speranze riposte su farmaci di recente formulazione: in commercio sarà presto disponibile il preparato denominato T20, mentre un vaccino denominato Aidsvax è stato oggetto di sperimentazione sull'uomo per uno studio i cui risultati verranno

pubblicati all'inizio del 2003. Resta purtroppo argomento comune ai precedenti meeting internazionali quello della inaccessibilità ai farmaci per gran parte dei paesi in via di sviluppo.

L'AIDS nel mondo

Poche zone della Terra sono state risparmiate dal virus HIV, che nelle diverse aree si è diffuso con modalità differenti. Ad esempio, in alcune zone del Nord America, la gran parte delle persone infettate è rappresentata da omosessuali e uomini bisessuali, o da tossicodipendenti che fanno uso di siringhe e dai loro partners sessuali. In Africa e in Asia, i soggetti colpiti sono eterosessuali che contraggono l'infezione attraverso frequenti contatti sessuali con diversi partners. Circa il 60% dei portatori di HIV, sia sieropositivi sia malati di AIDS, vive nell'Africa sub-sahariana, il 24 % in Asia, il 5% nell'America Latina, il 5% in Nord America, più del 3% in Europa e meno del 2% nei Caraibi (**figura 1**). Si pensa che la sindrome abbia avuto origine nell'Africa sub-sahariana⁷ per mutazione di un retrovirus animale, forse della scimmia, che nel XX secolo fu

trasmesso alla popolazione umana diventando poi una epidemia globale. La UNAIDS e il WHO stimano 25 milioni di morti dalla scoperta della sindrome, il che ne ha fatto una delle più terribili epidemie della storia. Nel solo 2005 sono stati stimati circa 3,1 milioni di morti di cui 570.000 bambini.

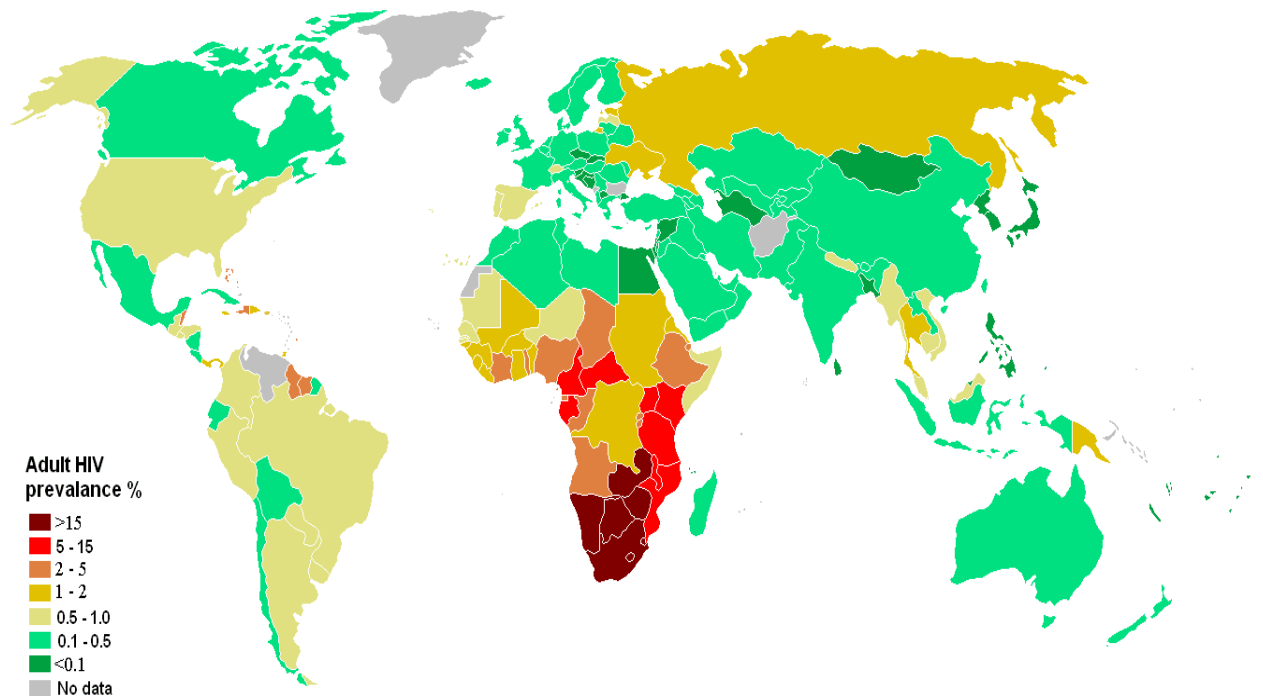


Figura 1. Estimated HIV/AIDS prevalence among young adults (15-49) by country as of 2008. Based on data from UNAIDS 2008 global report

Globalmente, un numero compreso tra 36,7 e 45,3 milioni di persone vive con l'HIV (fonte UNAIDS, 2005). Nel 2005, un numero compreso tra 4,3 e 6,6 milioni di persone è stato infettato e un numero compreso tra 2,8 e 3,6 milioni di persone è morto per l'AIDS, un incremento dal 2004 e il numero più alto dal 1981.

Il virus dell'HIV

L'HIV appartiene al genere dei Lentivirus, facente parte della famiglia dei Retroviridae. Le sue principali caratteristiche sono, come esprimono i nomi del genere e della famiglia di appartenenza, il comportamento da virus lento, cioè di virus che infetta l'organismo ospite, ma determina sintomi particolari solo dopo un certo periodo di tempo, e la presenza di un filamento di acido ribonucleico, RNA, come materiale genetico, caratteristica tipica di tutti i retrovirus^{6,8}.



Figura 2. Virus HIV. I retrovirus, cui appartiene il virus HIV, possiedono un patrimonio genetico costituito da un filamento di RNA. Nella foto, un virus HIV al microscopio elettronico. Corbis/Lester V. Bergman

Virus HIV-1 e HIV-2

Attualmente, si conoscono due ceppi di HIV: il virus HIV-1 ed il virus HIV-2. Il primo ha una diffusione su scala mondiale, mentre il secondo sembra localizzato nelle regioni centro-occidentali dell'Africa. L'HIV attacca in particolare i linfociti T-helper presenti nel sangue dell'organismo ospite, e in tal modo ne impoverisce le difese immunitarie. Nell'uomo causa una patologia nota

come sindrome da immunodeficienza acquisita (AIDS), che può manifestarsi anche a distanza di più di dieci anni dall'infezione. La scoperta dell'HIV come agente responsabile dell'AIDS, compiuta nel 1983 all'Istituto Pasteur di Parigi, è merito di Luc Montagnier. Altri gruppi di ricerca giunti allo stesso risultato poco tempo dopo sono quello di Robert Gallo, del National Cancer Institute di Bethesda, Maryland, e quello di Jay Levy, dell'Università della California a San Francisco.

Struttura del virione

Il virus dell'HIV, appartenente alla famiglia dei Retrovirus, genere Lentivirus, è un virus a RNA che, una volta penetrato nell'organismo, viene trascritto in DNA grazie all'enzima retrotrascrittasi (RT) e può così integrarsi nel genoma umano delle cellule bersaglio. Il provirus che in questo modo è venuto a formarsi è un intermedio replicativo necessario per la produzione di copie multiple di RNA e di proteine virali, che avviene utilizzando l'apparato trascrizionale dell'ospite. Queste componenti poi, una volta assemblate, danno

origine al virione completo che è così in grado di moltiplicarsi all'interno dell'organismo.

All'analisi con microscopio elettronico la particella virale si presenta di forma rotondeggiante con diametro compreso tra i 100 e i 120 nm. Esternamente si può distinguere un involucro pericapsidico detto envelope, che riveste il nucleocapside o core (**figura 2**).

L'envelope è costituito da un doppio strato fosfolipidico al quale sono ancorate le glicoproteine virali. Il doppio strato 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. Questo processo avviene generalmente a livello della superficie cellulare ma, nel caso dei macrofagi, i virioni maturano per lo più intracellularmente, gemmando attraverso il reticolo endoplasmatico.

Come conseguenza di questi fenomeni maturativi sull'envelope sono presenti perciò sia proteine virali che antigeni di superficie propri della cellula ospite (es. antigeni del complesso maggiore di istocompatibilità di classe I e II) e molecole di adesione (adesine).

Per quanto riguarda le proteine di origine virale, a livello della superficie esterna si ritrovano strutture di 9-10 nm formate da omotrimeri, i cui singoli

componenti sono a loro volta costituiti da due subunità proteiche associate, entrambe prodotte dal gene env.

Questo gene codifica per un precursore proteico di peso molecolare 160 kDa che, in seguito a scissione proteolitica, dà origine alla glicoproteina di membrana 120 (gp120), che sporge all'esterno della superficie, e alla glicoproteina 41 (gp41), che è collocata nello spessore della membrana (proteina transmembrana).

La glicoproteina gp120, di 515 residui aminoacidici, promuove la penetrazione del virus nella cellula legandosi a specifici recettori di superficie e il sito principale di interazione sul recettore è situato in un'ansa strutturale, che si trova nel primo dominio extracellulare Ig-simile del CD4⁹.

La glicoproteina gp41, di 345 residui aminoacidici, è responsabile sia dell'ancoraggio di gp120 al doppio strato lipidico, sia della fusione tra membrana cellulare e virale in seguito al legame con il recettore, grazie a un peptide fusogenico ricco di glicine e perciò idrofobico.

Nella porzione più interna del doppio strato si trova un guscio elettron-denso di 7 nm che ricorda la matrice presente nella maggior parte dei virus rivestiti di envelope. Questo strato è costituito dall'assemblaggio di diverse unità della proteina 17 (p17), la cui parte Nterminale viene miristilata in fase post-traduzionale. La miristilazione consiste nel legame di molecole di acido

miristico a residui di glicina posti all'N-terminale della proteina ed è fondamentale per ancorare p17 alla parete interna dell'involucro fosfolipidico e per garantire il corretto assemblaggio delle glicoproteine di tale involucro¹⁰⁻¹¹.

All'interno della struttura virale si trova il nucleocapside o "core" del virus, che presenta la tipica forma cilindrico-conica dei Lentivirus ed è costituito dalla proteina p24, codificata dal gene gag. Il dominio C-terminale di p24 riveste un ruolo importante per l'assemblaggio e permette la dimerizzazione della proteina, nonché la formazione di oligomeri gag, mentre il dominio N-terminale partecipa alla liberazione del genoma dal capsido, quindi è essenziale per l'infettività del virus.

All'interno del nucleocapside si trovano i componenti necessari alla replicazione del virus quali il genoma, costituito da due filamenti di RNA monocatenario a polarità positiva, molecole di tRNA necessarie per l'inizio della replicazione, proteine non strutturali prodotte da geni accessori e regolatori e i prodotti del gene pol, ovvero gli enzimi trascrittasi inversa, ribonucleasi H, integrasi e proteasi. Nel virione maturo il materiale genetico è complessato alle proteine basiche p7 e p9 prodotte, come p24 e p17, dal gene gag e necessarie per l'incapsidazione dell'RNA virale.

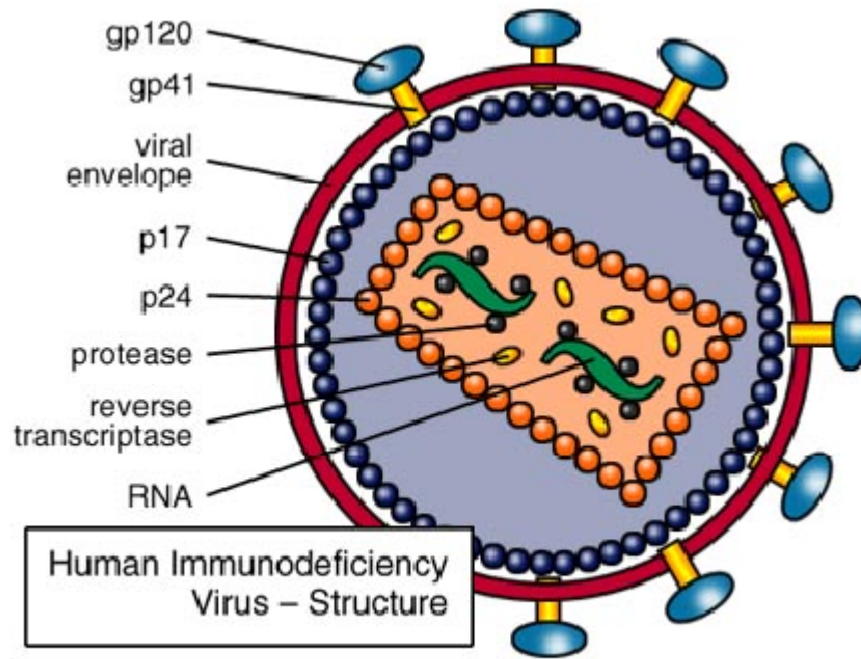


Figura 3. Struttura del virione di HIV-1

Il genoma di HIV

Il genoma virale misura circa 9,2 Kb ed è costituito da due filamenti di RNA monocatenario lineare a polarità positiva, uniti a livello delle estremità 5' attraverso legami idrogeno in corrispondenza di sequenze complementari¹²⁻¹³.

Le due estremità 3' libere sono poliadenilate (poli A), mentre all'estremo 5' è presente un "capping" (Cap), costituito da residui di guanosina fosforilata e metilata. A una distanza di circa un centinaio di nucleotidi dal 5', all'RNA è appaiata una molecola di tRNA di origine cellulare che funge da primer al momento della retrotrascrizione del genoma virale.

Ciascuna molecola di RNA genomico presenta, agli estremi, due sequenze di basi reiterate, 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).

Le sequenze U5 ed U3 sono retrotrascritte dalla trascrittasi inversa ad ambedue gli estremi della molecola di DNA provirale che risulta, in tal modo, più lunga rispetto alla molecola di RNA genomico. Si formano di conseguenza, ad entrambi gli estremi della molecola, sequenze identiche non codificanti e

altamente conservate (U3, R, U5) denominate LTR (long terminal repeat), che contengono il promoter e l'enhancer per la trascrizione del provirus.

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.

Funzionalmente dunque 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.

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.

Dall'estremo 5' è possibile individuare:

- TAR, sito di legame di Tat;
- Il sito di legame del primer, (residui 182-199), che fa da innescamento per la trascrizione inversa;

- Il segnale di incapsidamento ψ (residui 240-350), fondamentale per l'incorporazione dell'RNA nel virione;
- Il sito di dimerizzazione, (residui 248-271), che facilita l'incorporazione dei due RNA nel virione;
- Il maggior sito donatore di splicing, (residuo 290), utilizzato nella generazione di tutti gli mRNA subgenomici;
- La regione Gag-Pol, (residui 1631-1672), comprensiva di una sequenza eptanucleotidica e di una forcina, che promuovono lo spostamento del sito di lettura ribosomiale (-1), favorendo la traduzione di una poliproteina Gag-Pol con una frequenza del 5-10%;
- RRE (Rev Response Element), sito di legame di Rev;
- Siti accettori di splicing presenti in diverse regioni dell'RNA, che garantiscono la produzione di un elevato numero di prodotti di splicing (siti principali nei residui 5358 e 7971);
- Il segnale di poliadenilazione, (residui 9205-9210), che costituisce l'estremo 3'.

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.

Il genoma di HIV-1 possiede 9 geni fondamentali, di questi tre (gag, pol, env) sono geni strutturali tipici di tutti i Retrovirus, essenziali per codificare gli elementi necessari per la replicazione virale.

Vi sono poi geni regolatori (rev, tat) e geni accessori (nef, vpr, vpu e vif), che permettono ad HIV una più complessa interazione con la cellula ospite (figura 4 e tabella 1).

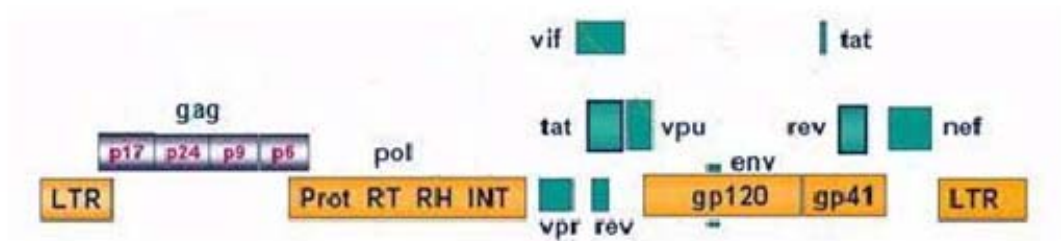


Figura 4. Mappa genetica di HIV-1

Tabella 1. (da: Immunologia Cellulare e Molecolare. Abbas – Lichtman – Pober)

LTR	Integrazione del DNA virale nel genoma della cellula ospite; sito di legame per i fattori di trascrizione della cellula ospite.
Gag	Proteine del nucleocapside e della matrice.
Pol	Trascrittasi inversa, proteasi, integrasi e ribonucleasi.
Env	Proteine dell'envelope virale (gp120 e gp40) che mediano il legame con CD4 e coi corecettori e la fusione della membrana del virus con quella cellulare.
Vif	Potenzia l'infettività delle particelle virali.
Vpr	Favorisce il trasporto nel nucleo del DNA virale; arresto nella fase G ₂ del ciclo cellulare.
Tat	Necessaria per l'estensione dei trascritti virali.
Rev	Favorisce l'esportazione dal nucleo di RNA virali non andati incontro a splicing.
Vpu	Riduce l'espressione di CD4 da parte della cellula e favorisce la liberazione del virus.
Nef	Riduce l'espressione di CD4 da parte della cellula e favorisce la liberazione del virus; riduce l'espressione di MHC di classe I da parte delle cellule infette.

Proteine strutturali di HIV

Analogamente a tutti i Retrovirus, il genoma di HIV contiene i tre geni strutturali 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. Dal taglio del precursore si forma così una proteina di 55 kDa (p55) e le tre proteine a funzione enzimatica associate al virione, rappresentate da una proteasi virus-specifica (p10 o PR), da una trascrittasi inversa (p51/66 o RT) e da un'integrasi (p32 o IN). La suddivisione di p55 porta alla formazione della proteina della matrice (p17 o MA), della proteina capsidica (p24 o CA), della proteina p15 che, nel virione in formazione, viene ulteriormente scissa nella proteina del nucleocapside (p9 o NC) e in p7, necessarie per un corretto incapsidamento dell'RNA. Il gene env viene tradotto in una proteina (p88) 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, gp120 e gp41.

Una volta sintetizzata, la poliproteina gp160 può interagire, a livello del reticolo endoplasmatico, con le molecole CD4 sintetizzate dalla cellula, ma tale legame viene contrastato dalla proteina virale Vpu, che induce la degradazione delle molecole CD4⁵. La glicoproteina gp160 viene quindi veicolata sulla superficie cellulare ed esposta all'esterno. Per impedire un'eventuale interazione tra essa e le molecole CD4 espresse in superficie interviene la proteina virale Nef, che promuove l'endocitosi e la degradazione delle molecole CD4 di superficie a seguito dell'infezione. Quindi due diversi messaggeri derivano dall'mRNA retrovirale: uno di lunghezza completa che porta alla traduzione delle poliproteine Gag e Pol, l'altro subgenomico, più breve, ottenuto grazie a un meccanismo di splicing, e che origina il prodotto del gene env. Per quanto riguarda le poliproteine Gag e Pol, la prima verrà prodotta dalla traduzione dell'mRNA dal codone d'inizio fino al primo codone di terminazione, il superamento del quale determina invece l'espressione di pol; a questo punto diversi meccanismi di traduzione possono intervenire in base agli schemi di lettura dei due geni. Quando essi sono in frame si ha la soppressione del codone di stop interposto tra le due sequenze, se invece gli schemi di lettura sono diversi o sovrapposti la poliproteina è prodotta attraverso un meccanismo di "scivolamento" ribosomiale. Questi meccanismi di antiterminazione sono però poco efficienti e perciò la produzione di Gag è circa 20 volte maggiore rispetto a

Gag-Pol. Oltre ai geni codificanti proteine strutturali, il genoma di HIV contiene anche 6 geni regolatori e accessori: i primi sono rappresentati da Tat e Rev, mentre i secondi da Nef, Vpr, Vif e Vpu. I prodotti di questi geni giocano un ruolo essenziale nel ciclo re plicativo virale e nella sintesi di macromolecole ad esso correlate (figura 5).

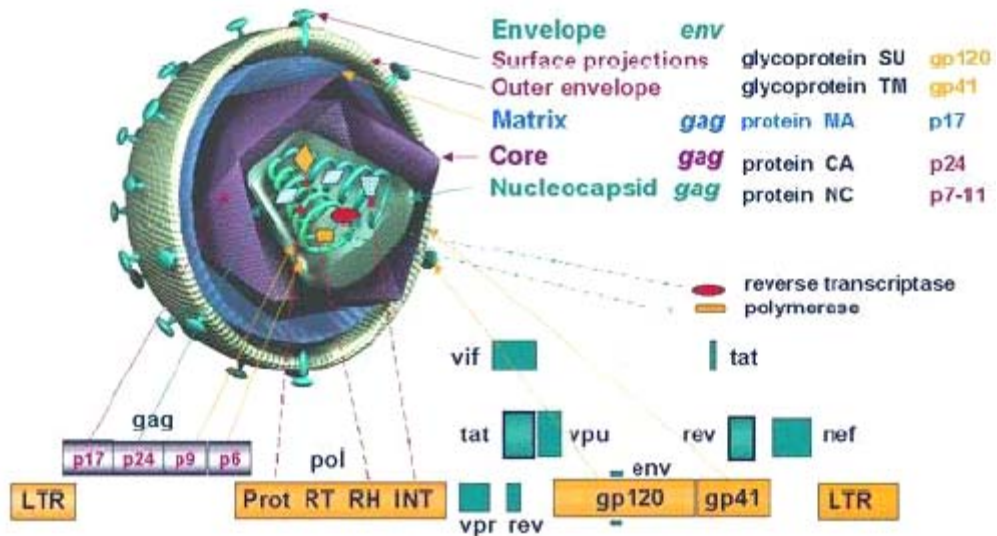


Figura 5. Struttura del virus: geni virali e le proteine da essi codificate

Poliproteina Pol

Il gene pol codifica per tre proteine virali a funzione enzimatica: la proteasi, la trascrittasi inversa e l'integrasi. La proteasi (PR o p10) è una proteina di 99 residui aminoacidici, che agisce in forma di dimero. Essa interviene nell'elaborazione delle proteine virali a partire dai precursori poliproteici e gioca perciò un ruolo fondamentale nella maturazione della particella virale. PR 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-Pol⁶. La proteina, incapsidata come parte del precursore Pol all'interno dei virioni al momento della loro fuoriuscita dalla cellula, attiva una serie di riarrangiamenti conformazionali della progenie ancora immatura e non infettante e promuove il taglio di precursori poliproteici, consentendo così la maturazione e l'infettività della particella virale. Il sito attivo dell'enzima, posto nell'interfaccia del dimero, è collocato a livello del residuo Asp25, esposto da ogni monomero ed essenziale per la funzione catalitica; inoltre è presente una sequenza altamente conservata caratteristica di tutte le aspartil-proteasi (Asp-Thr-Gly) e una struttura a cerniera importante per il legame dell'enzima al substrato. La Trascrittasi

inversa (RT) è un eterodimero con una subunità di 560 aminoacidi (p66) e una da 440 residui (p51), entrambe derivate dalla poliproteina Pol, contenenti ciascuna un dominio ad attività polimerasica costituito da quattro regioni. 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, grazie alla sua azione DNA polimerasica RNA dipendente. L'enzima viene incapsidato nella particella di HIV per poter svolgere la sua azione una volta avvenuta la liberazione del genoma virale dal capsido, fenomeno denominato uncoating.

La retrotrascrittasi, nella subunità p66, possiede anche un dominio ad attività RNasica (RNasi H), che, durante la replicazione, degrada l'RNA dell'ibrido DNA-RNA⁷⁻¹³, inducendo così la sintesi di un doppio filamento di DNA ed il distacco del tRNA dal filamento neoformato. Nonostante le sequenze aminoacidiche delle due subunità appaiano simili, solo sulla p66 è presente il sito catalitico, caratterizzato da tre residui altamente conservati (Asp 110, Asp185, Asp186), mentre la p51 si presenta con una conformazione chiusa, funzionalmente inattiva¹⁴.

Il meccanismo enzimatico in cui è coinvolta la trascrittasi inversa ha inizio all'interno della particella virale, quando questa non ha ancora iniziato la sua azione infettiva verso la cellula bersaglio. A questo livello avviene l'annealing, ovvero l'appaiamento del tRNA che fa da primer, di origine cellulare e posto

all'interno del virione durante l'assemblaggio, col genoma virale, in particolare con la sequenza PBS (Primer Binding Site), posta vicino all'estremo 5'. Il tRNA contatta così sia l'interfaccia del dimero della retrotrascrittasi che la regione C-terminale di p66. Qualunque tRNA può essere utilizzato dall'RT come primer, ma si è visto come la retrotrascrizione sia più efficiente qualora sia presente un tRNA caricato con l'aminoacido lisina (tRNA^{Lys}) oltre ad un'appropriata concentrazione di ioni Mg²⁺.

Le reazioni successive avvengono nella cellula ospite, a livello citoplasmatico, dopo la liberazione del genoma virale. Caratteristica tipica della retrotrascrittasi virale è la sua scarsa fedeltà replicativa, essendo priva di un'attività esonucleasica di controllo in direzione 3'-----5'; questa caratteristica dell'enzima porta alla formazione di circa 10 errori per ciclo di sintesi, originando così un'alta frequenza di variabilità nucleotidica fra i diversi ceppi. Da ciò risulta evidente come sia quindi difficile elaborare una strategia terapeutica efficace contro questo virus, capace di modificare con tale rapidità le proprie caratteristiche antigeniche ed enzimatiche. L'endonucleasi-integrasi (IN o p32) è una proteina di 288 aminoacidi che agisce sotto forma di tetramero a livello del nucleo cellulare, dove catalizza una serie di passaggi che portano all'integrazione del DNA provirale nel genoma della cellula ospite¹⁵.

Strutturalmente ogni monomero è costituito da tre domini caratterizzati da una specifica struttura: il dominio N-terminale, quello catalitico e il C-terminale. Il primo dominio, compreso tra gli aminoacidi 1 e 50, è coinvolto nella dimerizzazione e contiene un sito di legame per lo zinco coordinato con due cisterne e due istidine, inoltre contiene una struttura HLH (helix-loop-helix) analoga a quella presente nelle proteine leganti il DNA.

Il dominio catalitico, residui 50-212, è caratterizzato dalla presenza di un motivo conservato, il quale lega Mg^{2+} o Mn^{2+} , ioni metallici necessari perché avvenga l'integrazione.

L'ultimo dominio, compreso tra i residui 220 e 270 e strutturalmente simile al dominio SH3, è coinvolto sia nella dimerizzazione, sia nella formazione di legami aspecifici con il DNA; contiene infatti un solco a forma di doppia sella che stabilisce un legame con i filamenti di DNA tramite la Lys264 contenuta nella sequenza del dominio. All'interno delle cellule infette si possono osservare tre forme di DNA virale: DNA lineare delimitato da due LTR, DNA circolare con una LTR e DNA circolare con due LTR. Solo il DNA lineare è substrato dell'enzima integrasi. L'integrazione può avvenire in diversi siti essenzialmente casuali, benché si sia evidenziata una preferenza per le sequenze di DNA particolarmente condensate, contenute nella struttura nucleosomica. Il provirus integrato presenta alcune differenze rispetto alla forma non integrata, infatti ha

perso alcuni nucleotidi alle estremità e le sequenze virali sono affiancate da brevi sequenze ripetute del sito bersaglio dell'ospite.

È possibile distinguere diverse fasi nel processo di integrazione: la prima consiste nel riconoscimento delle estremità del DNA virale a livello dei siti AAT e nella rimozione, da parte dell'enzima, di due paia di basi dalle estremità 3' di ogni filamento di DNA del virus, esponendo così il dinucleotide CA, altamente conservato tra i retrovirus. Nella seconda fase tali estremità si legano covalentemente con le estremità 5' del DNA cellulare, a sua volta clivato in corrispondenza del sito bersaglio di integrazione. Nel terzo passaggio si ha la riparazione, pare ad opera di enzimi cellulari, delle giunzioni tra DNA virale integrato ed il DNA genomico della cellula ospite e si generano così le ripetizioni dirette di 5 paia di basi che affiancano le sequenze virali. Nel meccanismo di integrazione sembrano essere coinvolte anche alcune proteine cellulari, ad esempio Ini1, che influisce sull'efficienza della trascrizione e HMG I, associata alla formazione di un complesso di pre-integrazione.

Poliproteina Env

Il gene env codifica per una proteina di 160 kDa (gp160), che viene scissa nelle glicoproteine gp120 e gp41. Gp120 o SU (surface) è una proteina di 515 residui aminoacidici, 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. CD4 è presente su linfociti T e monociti/macrofagi ed è costituito da quattro domini extracellulari Ig-simili. Il legame gp120-CD4 non è però sufficiente a consentire l'ingresso del virus nella cellula, è infatti richiesta anche la presenza di specifici recettori per le chemochine (una famiglia di molecole coinvolte in svariate attività biologiche, ad esempio la chemiotassi cellulare), che svolgono la funzione di corecettori virali⁹. Sulla base della specificità di queste molecole HIV può essere suddiviso in due classi: alla prima appartengono i virus a tropismo linfocitario sincizio-inducenti, che riconoscono come corecettore la proteina CXCR4, alla seconda i ceppi a tropismo macrofagico non inducenti la formazione di sincizi, che riconoscono il corecettore CCR5.

È il loop V3 di gp120 a recare determinanti che specificano selettivamente l'interazione con i corecettori cellulari, fornendo la chiave molecolare del tropismo verso linfociti T o macrofagi. Variazioni conformazionali della glicoproteina gp120 in seguito al legame con CD4, con la conseguente esposizione di domini del peptide fusogeno situati all'N-terminale di gp41, regolano inizialmente il processo di fusione di HIV con la membrana cellulare. L'analisi delle sequenze aminoacidiche di gp120 ha rivelato la presenza di cinque regioni discontinue, contenenti residui altamente variabili (V1-V5), distribuite in regioni altamente conservate. Il dominio V3, esposto come conseguenza del legame gp120-CD4, è il principale determinante di neutralizzazione e sembra rivestire un ruolo primario nelle fasi di infezione successive al legame; anticorpi diretti contro il dominio, infatti, bloccano la fusione tra virus e cellula senza inibire l'adesione a CD4.

Perciò in gp120 sono presenti due zone ben distinte, la prima responsabile dell'ancoraggio al recettore CD4 (residui 400-430), la seconda delle modificazioni della molecola che portano all'esposizione del peptide fusogeno (residui 313-337). Gp120 poi interviene direttamente nella patogenesi della malattia inducendo, dopo interazione con CD4 delle cellule sensibili, una serie di segnali negativi per la sopravvivenza cellulare e dando inizio al processo di morte cellulare programmata o apoptosi, anche in assenza di infezione da parte del

virus. Esistono inoltre ulteriori modalità di legame che permettono ad HIV di penetrare nelle cellule, ad esempio il legame di virus rivestiti di anticorpi ai recettori per il frammento Fc delle Ig (ciò spiega l'aumentata infettività virale mediata da anticorpi), e l'elevata affinità di gp120 per il glicolipide galattosio-ceramide, presente sulla superficie di linee cellulari gliali e di neuroblastoma. Gp41 o TM è una proteina di 345 residui aminoacidici situata nella membrana virale, con l'estremo COOH all'interno e l' NH_2 all'esterno del virione. 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 NH_2 e viene esposta in seguito al legame di gp120 al recettore CD4^{16} .

Poliproteina Gag

Il gene gag codifica per quattro proteine originate dal taglio proteolitico del precursore p55: La proteina della matrice, la proteina capsidica, la proteina del nucleocapside e p6.

La proteina della matrice (MA o p17), componente N-terminale della poliproteina Gag, si posiziona sul lato interno della membrana cellulare nelle zone in cui il virione fuoriuscirà dalla cellula, andando così a formare la matrice della particella virale.

Essa riveste un ruolo importante per il trasporto dei precursori gag e gag-pol verso i siti di assemblaggio virale a livello della membrana cellulare. La funzione della proteina è resa possibile da due proprietà strutturali: la presenza di un gruppo miristilato N-terminale e di residui basici localizzati a livello dei primi 50 aminoacidi. Dal punto di vista strutturale ma si presenta in una forma trimerica funzionalmente attiva, in cui i residui basici sono a livello della superficie del trimero¹¹. Il corretto posizionamento della proteina sulla membrana citoplasmatica si basa innanzitutto sull'inserimento di ciascun gruppo miristilato nel doppio strato lipidico, seguito da interazioni tra i residui basici e la parte idrofila dei fosfolipidi di membrana. Contribuisce, inoltre, all'inserimento delle

glicoproteine derivate da env all'interno della particella virale; queste proteine sono caratterizzate da lunghe code citoplasmatiche, che vengono inserite nelle cavità originate nella configurazione spaziale della proteina trimerica. È stato infine osservato un coinvolgimento della proteina della matrice nell'induzione dell'infezione di cellule non proliferanti, in particolare i macrofagi, grazie alla presenza di un sito NLS (Nuclear Localization Signal), che facilita il trasporto del complesso ribonucleico virale liberato dal capsido verso il nucleo della cellula infettata. La proteina del capsido (CA o p24), presente in circa 2000 copie in ogni struttura virionica, costituisce il "core" del virus. Il dominio C-terminale della proteina (residui 152-231) interviene nell'assemblaggio del virus, nella dimerizzazione della proteina stessa e nella formazione di oligomeri Gag. È costituito da una struttura distesa seguita da quattro α -eliche e contiene una sequenza compatta detta MHR (Major Homology Region), in cui si trovano venti aminoacidi molto conservati, stabilizzata da legami idrogeno tra i quattro residui a minor variabilità¹². È la regione MHR che, interagendo con Gag, permette l'ingresso di Gag-Pol nel virione in formazione. Il dominio N-terminale (residui 1-151) è invece essenziale per l'infettività poiché, grazie all'interazione con una chaperonina cellulare, la ciclofillina A, promuove l'incorporazione della proteina nel virione e partecipa all'uncoating virale. La proteina del nucleocapsido (NC o p9), formata da 55 residui aminoacidici, contiene due

domini zinc fingers fiancheggiati da residui basici che permettono che essa si complessi con il genoma del virus. La funzione primaria di NC è quella di portare l'acido nucleico completo all'interno della struttura virionica in formazione, legandosi tramite le strutture zinc fingers al segnale di incapsidamento (ψ packaging signal) dell'RNA, strutturalmente costituito da tre strutture a forcina localizzate in prossimità del sito donatore di splicing maggiore. La più nota delle tre è la forcina che ha al suo interno la regione del genoma responsabile della dimerizzazione delle due copie genomiche presenti nel virione maturo. Poiché p9 è una proteina basica, riveste l'RNA genomico nel core della particella, proteggendolo in questo modo dalle nucleasi. Inoltre, come fattore chaperon, partecipa all'annealing del tRNA-primer, allo srotolamento delle strutture secondarie dell'RNA, alle reazioni di scambio dei filamenti durante la retrotrascrizione e nell' induzione dell'integrazione. La proteina p6, costituita da 51 aminoacidi C-terminali di Gag, è importante per l'incorporazione di Vpr durante l'assemblaggio virale. La proteina contribuisce inoltre all'efficiente rilascio del virione dalla cellula ospite.

Proteine regolatrici ed accessorie di HIV

Il gene Tat (Trans-Activator of Transcription) è costituito da due esoni e codifica una proteina di 86 aminoacidi, del peso di 14-15 kDa, la cui forma funzionante deriva da splicing multipli dell'mRNA. Tat rappresenta uno dei primi prodotti del virus poiché svolge un ruolo essenziale nel trans-attivare la trascrizione delle proteine virus-specifiche agendo sul complesso LTR 5' e consente quindi di incrementare di circa 100 volte la produzione di mRNA virale¹⁷⁻¹⁸.

Essa può:

- legarsi, tramite un dominio ricco in arginine, ad una struttura a loop alla sequenza TAR (tat responsive region) presente al 5' degli mRNA virali nascenti e facilitare il reclutamento di fattori trascrizionali cellulari e di adattatori trascrizionali, potenziando così la funzione del complesso trascrizionale virale (meccanismo TAR-dipendente).

- interagire attraverso un loop autocrino con recettori di membrana della cellula infetta, innescando così una cascata di “signalling” al nucleo che si traduce in un aumentata attivazione dei fattori trascrizionali disponibili nella cellula (c-Fos, Creb), con conseguente incremento della trascrizione del provirus (meccanismo TAR-indipendente).

Il legame di Tat alla struttura ribonucleica prevede una prima interazione tra la sequenza ricca di arginine ed una guanina presente nel solco maggiore dell'RNA, seguita dalla stabilizzazione del complesso sia tramite contatti addizionali coinvolgenti i gruppi fosfato presenti nella struttura dell'acido nucleico, sia attraverso la formazione di un triplice appaiamento U-A:U tra una base localizzata in una protuberanza di TAR e una coppia di basi presenti più a monte.

Tat è necessaria per stabilizzare i processi trascrittivi, conferisce “processività” alla RNA polimerasi II cellulare e può anche influire sul tasso di iniziazione della trascrizione. Pare che essa arruoli o attivi chinasi che fosforilano il dominio C-terminale dell'RNA polimerasi II, processo che favorisce la conversione di tale enzima da iniziatore a fattore di elongazione, impedendo così una terminazione prematura del trascritto. Tat può poi essere eliminata

all'esterno della cellula infetta e andare a interagire con le cellule circostanti non infette (azione paracrina), le quali la internalizzano per endocitosi. Tat può perciò agire transattivando la trascrizione di numerosi geni cellulari produttori di fattori di crescita e citochine (TNF, IL6).

Ricerche recenti hanno inoltre dimostrato che cellule transfettate col gene tat risultano protette dall'apoptosi e hanno quindi un'aumentata capacità di sopravvivenza e proliferazione. In realtà pare che questa proprietà di Tat sia funzione della sua concentrazione: a concentrazioni fisiologiche svolge l'azione protettiva, mentre a concentrazione elevate ha effetto opposto, inducendo la comparsa dell'apoptosi. Il gene Rev (Regulator of Virion Expression), costituito da due esoni sovrapposti a quelli di tat, codifica per una proteina di 116 residui aminoacidici (p19) a localizzazione prevalentemente nucleare, che gioca un ruolo cruciale nell'infezione virale per il suo ruolo di regolatore "differenziale". Essa consente infatti la transizione dalla produzione di trascritti primari sottoposti a più eventi di splicing (che codificano le proteine regolatorie e accessorie Tat, Rev e Nef), a trascritti prodotti da un singolo splicing e codificanti proteine strutturali. In questo modo il gene controlla il passaggio da un'infezione quiescente ad una crescita virale attiva¹⁹.

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, sull'inibizione dello splicing prima dell'esportazione 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) si trova nella regione 3'LTR e codifica per una proteina di 27 kDa (p27) miristilata all'N-terminale e a localizzazione prevalentemente citoplasmatica, che è la prima proteina ad accumularsi nelle cellule successivamente all'infezione con HIV. Si pensava che svolgesse un'azione di regolazione negativa della trascrizione virale, che veniva così rallentata. Si è successivamente visto che essa possiede diverse attività: come Vpu maschera l'espressione di molecole MHC di classe I e inoltre riduce i livelli di espressione di CD4 alla superficie della cellula, stimolando l'endocitosi di CD4 e la loro distruzione lisosomiale.

Nef poi si associa con diverse chinasi cellulari nei linfociti T helper ed aumenta l'infettività del virus. Se Nef è presente i virioni prodotti sono dieci volte più infettanti. Il gene Vpr (Viral Protein R) , localizzato tra i geni pol ed env e parzialmente embricato con il gene vif e con l'esone iniziale di tat, 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 a specifiche sequenze di traslocazione nucleare dette NLS (Nuclear Localization Signal). Esse sono particolarmente importanti per la localizzazione nucleare in cellule quiescenti come i macrofagi, poiché consentono il trasporto anche in assenza di breakdown mitotico del nucleo. Gli NLS sono costituiti da due α -eliche N-terminali, responsabili della localizzazione della proteina a livello dei pori nucleari²⁰. 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. L'intervento di Vpr a questo livello avviene nella fase precedente la defosforilazione di una chinasi ciclina-dipendente necessaria per iniziare il processo mitotico. Per l'arresto in G₂ sono di fondamentale importanza gli aminoacidi localizzati nella regione C-terminale della proteina virale e si è inoltre osservato come diverse proteine cellulari interagiscano con essa, ad esempio la subunità regolatoria della fosfatasi PP2A (Protein Phosphatase 2A), che regola il passaggio dalla fase G₂ alla mitosi.

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 liberi²¹. Essa viene inoltre regolata tramite modifiche post-trascrizionali; si è visto infatti come la sua funzionalità sia alterata da mutazioni che interessano uno dei tre siti di fosforilazione del residuo Ser144.

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. Inoltre Vif pare sia coinvolta anche nell'adsorbimento e penetrazione nella cellula ospite, infatti ceppi virali di laboratorio, privi del gene Vif, sono ancora in grado di replicarsi e produrre virioni normali, incapaci però di infettare cellule sensibili. 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 vif²².

Il gene Vpu (Viral Protein U), codifica per una proteina di membrana di 81 aminoacidi, del peso di 16 kDa (p16), con un dominio citoplasmatico C-terminale e uno idrofobico terminale di 24 residui. Anche questa proteina viene regolata attraverso modificazioni post-trascrizionali, va infatti incontro a una fosforilazione a livello dei residui Ser52 e Ser56. Funzionalmente la proteina

interviene nell'assemblaggio, nella maturazione del virione e nella liberazione della progenie virale. Per quanto riguarda il processo maturativo, p16 promuove la degradazione di CD4 responsabili della ritenzione, a livello del reticolo endoplasmatico, delle proteine Env neosintetizzate, promuovendo così la loro esposizione sulla membrana cellulare. Il meccanismo alla base della degradazione pare coinvolgere il proteasoma cellulare, poiché un inibitore di quest'ultimo blocca l'intero processo. 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. Vpu può infine stimolare, attraverso il suo dominio N-terminale, il rilascio non specifico delle particelle virioniche, agendo sotto forma di canale ionico del potassio.

Infezione e ciclo vitale di HIV

Il virus HIV attacca specificamente alcuni tipi di cellule umane: i macrofagi; un sottogruppo di linfociti T-helper caratterizzati dalla presenza, sulla superficie esterna della membrana plasmatica, di recettori proteici indicati con le sigle CD4 e CCR5. Questi linfociti vengono perciò chiamati linfociti T-CD4+.

Dal rivestimento esterno del virus sporgono due tipi di glicoproteine, le gp120 e le gp41. La gp120 viene riconosciuta e legata dai recettori CD4; questo fenomeno induce una modificazione della struttura della gp120, che si lega anche al recettore CCR5. La formazione di tale complesso a sua volta determina uno scatto della glicoproteina gp41 verso la membrana plasmatica della cellula ospite e, dunque, l'avvio dell'infezione da parte del virus. L'HIV inietta il suo patrimonio genetico, ovvero i due filamenti di acido ribonucleico (RNA), e i suoi enzimi (trascrittasi inversa, proteasi e integrasi), nel citoplasma della cellula ospite.

La trascrittasi inversa dà inizio alla sintesi di un filamento di acido desossiribonucleico (DNA) complementare a ciascun filamento di RNA; si forma dunque un doppio filamento ibrido di DNA ed RNA. Infine, l'enzima degrada

la porzione di RNA e completa la sintesi di una molecola di DNA a doppio filamento.

L'enzima integrasi determina l'integrazione del DNA virale entro il DNA della cellula ospite (formazione del cosiddetto provirus); questo patrimonio genetico ibrido, sfruttando gli organuli della cellula ospite, dirige la sintesi di nuove proteine e componenti virali. Le proteine virali neosintetizzate si trovano in una forma inattiva; per azione dell'enzima proteasi, vengono tagliate in modo da convertirsi nella forma attiva. Quando i virus neoformati fuoriescono dalla cellula ospite, rimangono avviluppati da una porzione della membrana plasmatica, che costituisce il rivestimento esterno al capsido proteico, tipico di questi retrovirus. La cellula ospite, ormai degradata, muore.

La replicazione del virus HIV in alcuni tipi di cellule del sistema immunitario, illustrata di seguito segue un andamento caratteristico. Durante le prime due settimane dal contagio, l'HIV prolifera diffondendosi nei linfonodi (infezione primaria), generando sintomi analoghi all'influenza (sindrome acuta da HIV); questa fase acuta di infezione primaria persiste per circa tre mesi e causa una sensibile distruzione dei linfociti TCD4. Dopo sei mesi, nell'individuo si stabilisce un equilibrio tra i virus attivi e quelli che vengono neutralizzati dalle difese naturali (set point). Segue una fase di latenza, asintomatica, che prosegue per 8-10 anni, in cui il livello di virus nel sangue si abbassa; infine, compaiono

diverse malattie, che costituiscono il quadro clinico dell'AIDS conclamata, e per le quali la prognosi è nella maggior parte dei casi infausta.

Le fasi principali del ciclo vitale di HIV sono riportate in **figura 6**.

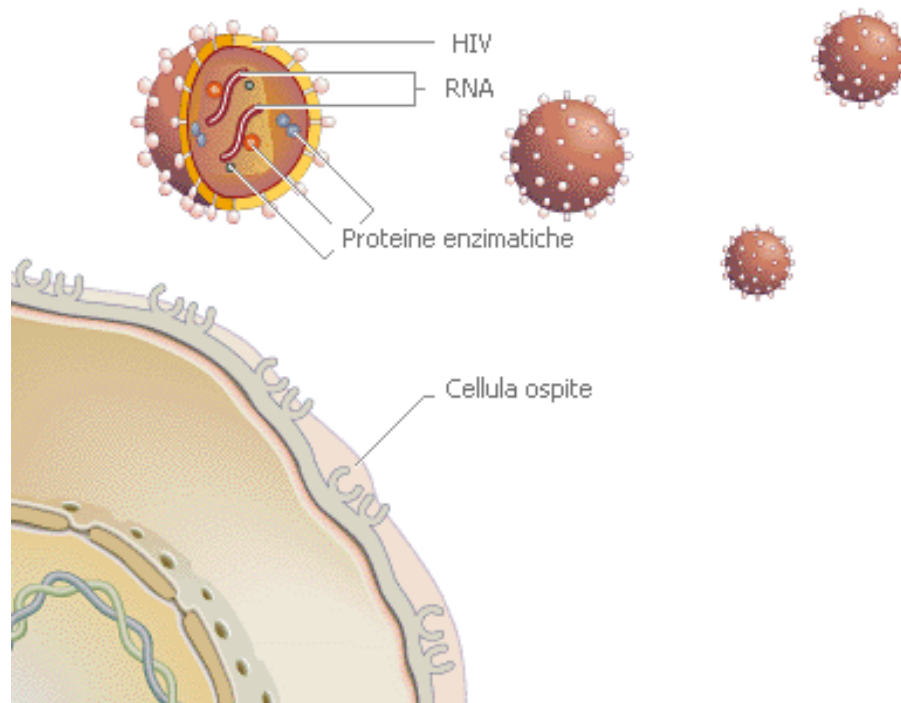


Figura 6a. Virus dell'HIV. Il virus HIV (Virus dell'Immunodeficienza Umana) è costituito da un involucro proteico (capside) che racchiude due eliche di RNA (acido ribonucleico), il patrimonio genetico di questo retrovirus. All'esterno vi è una membrana lipoproteica, nella quale si trovano inserite molecole glicoproteiche (gp120 e gp41). L'HIV possiede inoltre un corredo enzimatico che comprende la trascrittasi inversa, la proteasi e l'integrasi. Come tutti i virus, non può riprodursi al di fuori di una cellula; in particolare, l'HIV si replica sfruttando le cellule del sistema immunitario umano portatrici di recettori proteici detti CD4, quali un gruppo di linfociti T e di macrofagi.

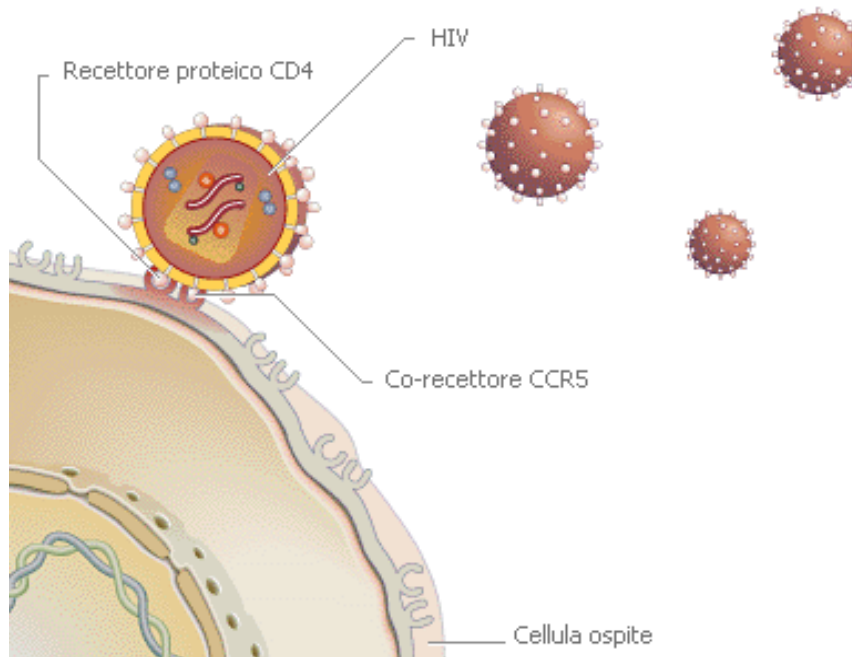


Figura 6b. Il legame dell'HIV al linfocita.

Il virus HIV si lega mediante una glicoproteina gp120 a un recettore CD4 che sporge dalla membrana del linfocita. Questa interazione rende la gp120 capace di legarsi al co-recettore proteico CCR5, anch'esso presente sulla superficie della cellula ospite; come conseguenza, le glicoproteine gp41 cambiano conformazione e si infiggono nello spessore della membrana.

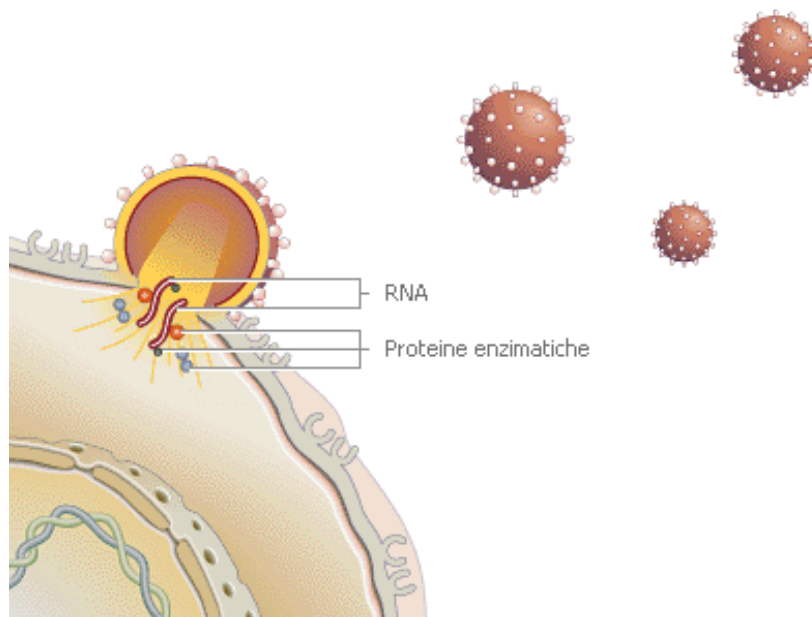


Figura 6c. Infezione del linfocita.

L'HIV inietta nel citoplasma del linfocita il suo patrimonio genetico e il corredo di enzimi che serviranno per la replicazione del virus.

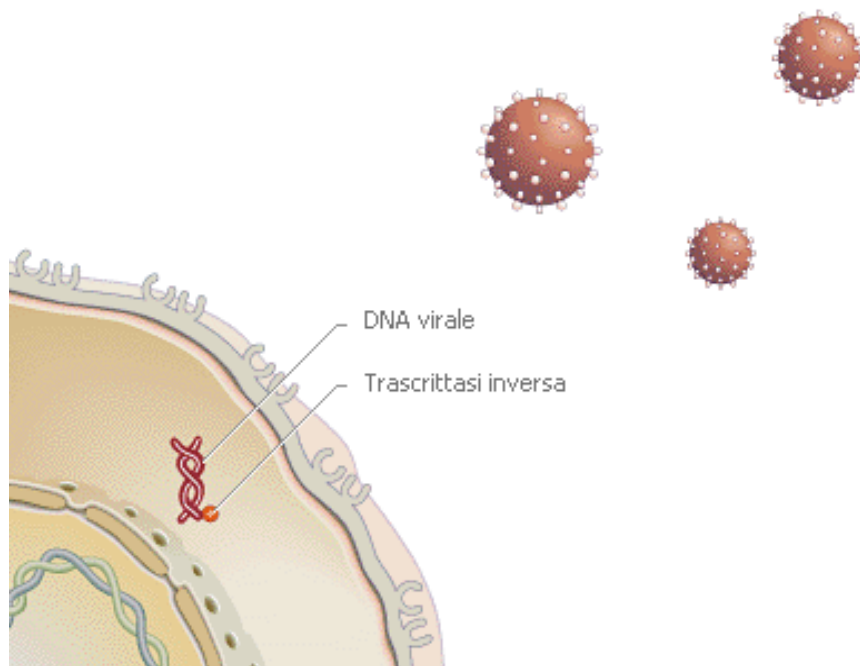


Figura 6d. Conversione dell'RNA virale in DNA

L'elaborazione di nuove particelle virali richiede che l'RNA venga prima convertito in DNA, cioè nell'acido nucleico sul quale si basa il processo di sintesi delle proteine. L'enzima virale trascrittasi inversa permette la conversione di un singolo filamento di RNA in una doppia elica di DNA.

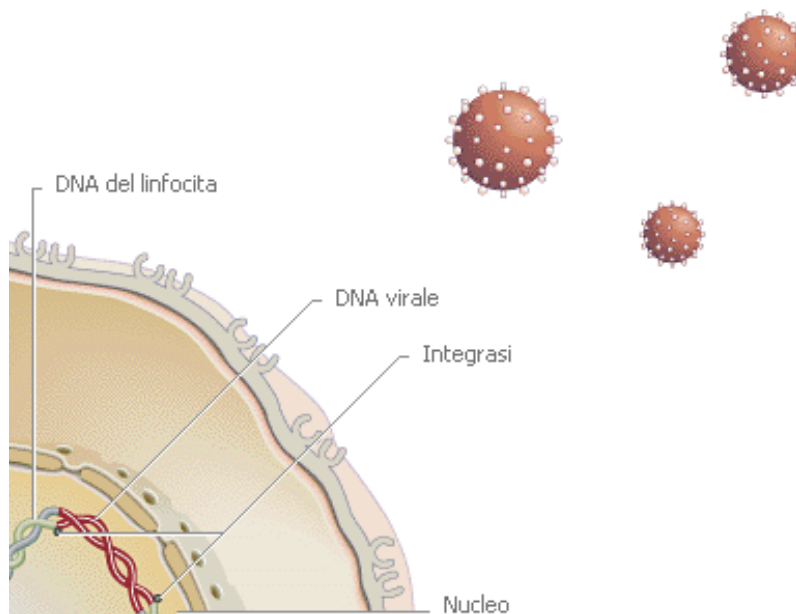


Figura 6e. Integrazione del DNA virale nel DNA del linfocita.

A livello del nucleo, l'enzima integrasi determina l'apertura del DNA del linfocita, e il successivo legame alle due estremità del DNA virale, che diviene così parte integrante del patrimonio genetico cellulare.

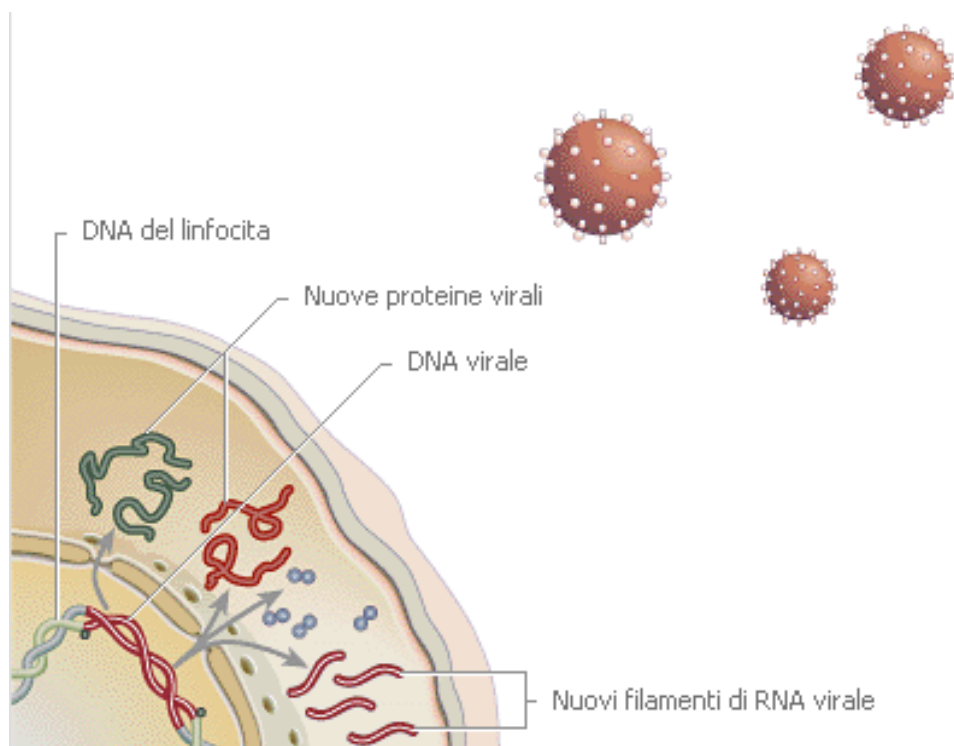


Figura 6f. Trascrizione

Il DNA cellulare compie normalmente il processo di trascrizione, mediante il quale dal DNA si formano molecole di RNA; poiché contiene DNA virale, vengono sintetizzate molecole di RNA dell'HIV, che dal nucleo si trasferiscono al citoplasma. L'RNA dirige il processo di traduzione, che avviene a livello dei ribosomi del linfocita e porta alla sintesi di nuove proteine virali.

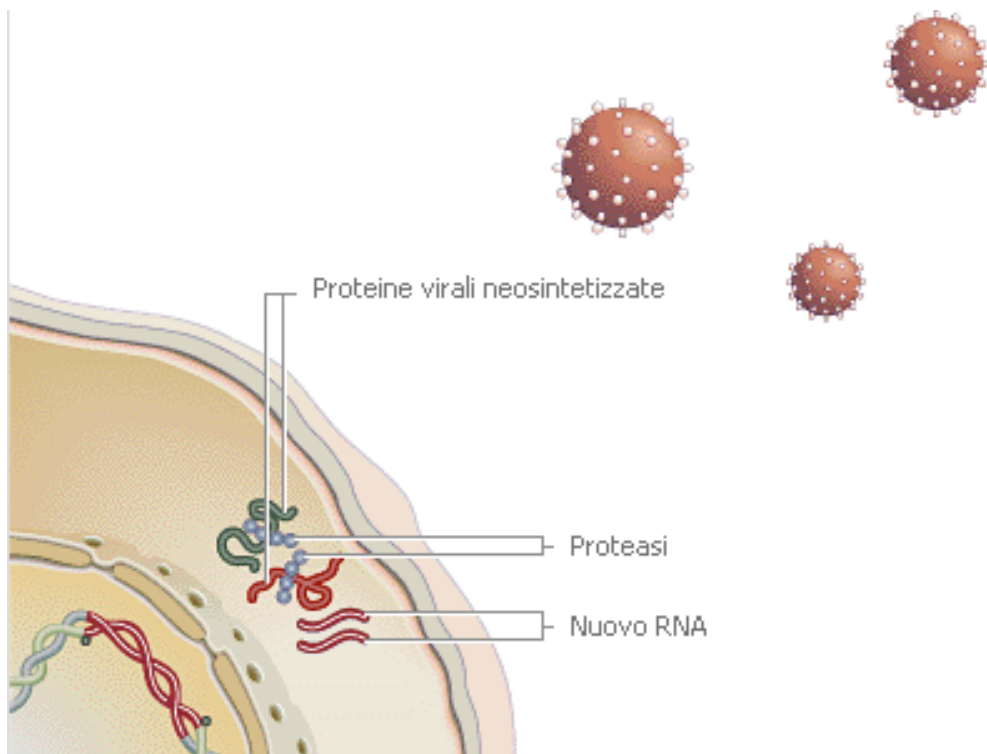


Figura 6g. Assemblaggio.

Il terzo tipo di enzima virale, la proteasi, taglia i filamenti proteici neosintetizzati, in catene proteiche più corte, che corrispondono a nuove particelle enzimatiche del virus e alle proteine che costituiranno nuovi capsidi.

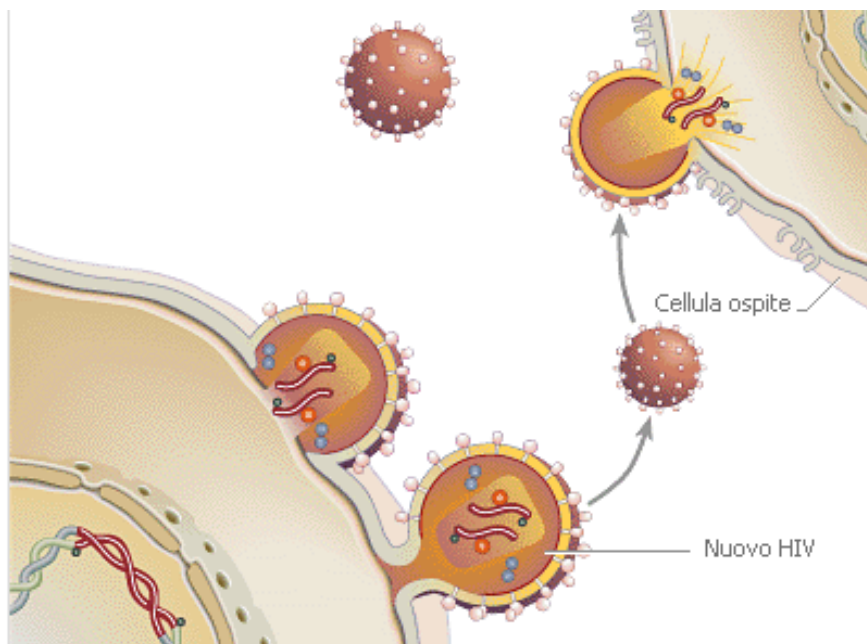


Figura 6h. Gemmazione.

Le proteine virali e i nuovi filamenti di RNA vengono inseriti nei capsidi. Si formano così particelle virali che, fuoriuscendo dal linfocita (gemmazione), restano avvolte dalla membrana cellulare e assumono la struttura caratteristica dell'HIV. La cellula ospite viene distrutta.

Infezione da HIV dei linfociti

La specificità del riconoscimento dei linfociti da parte del virus è determinata dalla presenza sulla superficie dei primi di un particolare tipo di proteina, la molecola CD4, prodotta dal sistema immunitario. La gp120 virale riconosce specificamente la molecola CD4 dell'ospite e si lega a essa, permettendo, così, la fusione del virus con la membrana cellulare dell'ospite e l'introduzione del materiale genetico virale nella cellula umana. Avvenuta l'infezione, si verifica la trascrizione inversa, processo mediante il quale, a livello del citoplasma cellulare, avviene la trascrizione dell'RNA del virus in una doppia elica di DNA; questa viene, poi, trasportata nel nucleo della cellula ospite e integrata nel DNA cromosomico umano. Sfruttando le strutture e i processi metabolici della cellula ospite, il DNA virale viene poi trascritto per formare nuove molecole di RNA virale e molecole di RNA messaggero (mRNA) virale. Questo presiede, a livello dei ribosomi della cellula ospite, alla sintesi di proteine virali. Queste vengono assemblate in modo da formare nuovi capsidi. Le particelle di HIV neoformate vengono liberate per gemmazione dalla superficie cellulare: si forma cioè una protuberanza sulla superficie cellulare, che infine si

rompe, liberando le particelle virali, ognuna delle quali resta avvolta in una porzione di membrana.

La progressiva distruzione dei linfociti T determina l'indebolimento del sistema immunitario dell'organismo. A causa di ciò, l'individuo può venire colpito dalle infezioni e dal cancro, tipici del quadro sintomatologico dell'AIDS (figura 7).



Figura 7. Linfocita T infettato dall'HIV.

Il virus HIV è la causa della sindrome da immunodeficienza acquisita (AIDS): agendo sulle proteine CD4 della membrana cellulare, esso penetra nei linfociti T responsabili delle difese immunitarie dell'organismo, esponendo l'individuo infetto a ogni tipo di malattia. Phototake NYC/Institut Pasteur/CNRI

Patogenesi di HIV

L'infezione da HIV evolve, nella maggioranza dei casi, secondo un andamento che comprende tre fasi successive: l'infezione acuta, una fase di latenza clinica, la sindrome da immunodeficienza acquisita. L'evoluzione del decorso dell'infezione può essere molto variabile, potendo essere influenzata da svariati fattori, primo tra tutti l'impiego di un'adeguata terapia antiretrovirale.

Nella maggioranza dei casi di pazienti senza presidi terapeutici tutto il processo si svolge nell'arco di 7-10 anni, traducendosi in un quadro patologico drammatico e che conduce alla morte del soggetto. A questo schema fanno eccezione i "Long-term nonprogressor", che non mostrano sintomi per molti anni e che non hanno una perdita significativa di cellule CD4+ (**figura 8**).

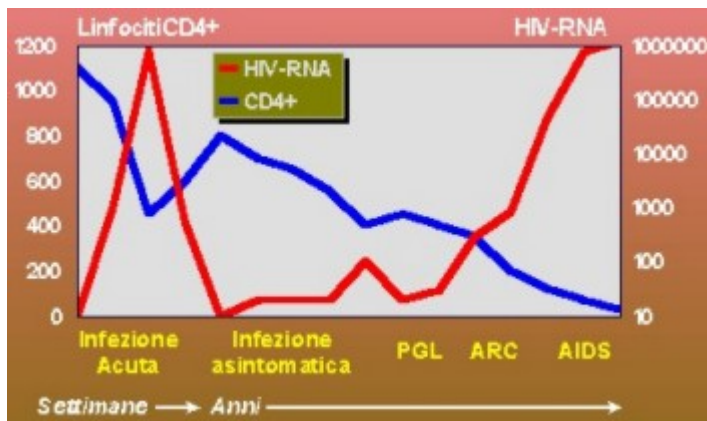


Figura 8 Andamento della carica virale e dei CD4⁺ nel corso delle diverse fasi della malattia (www.paginemediche.it)

Infezione acuta o primaria

Il periodo immediatamente successivo al contagio, viene definito come fase iniziale dell'infezione o infezione acuta.

Dal punto di vista virologico si accompagna ad un elevato titolo di HIV nel sangue del paziente (viremia >10⁶ virioni/ml di plasma) e ad un elevato tasso di replicazione che permette la disseminazione del virus a livello degli organi linfoidi (soprattutto nei linfonodi, ma anche in milza, tonsille e tessuto linfoide, che sono i tessuti più ricchi in linfociti CD4⁺). Questa disseminazione avviene

ancor prima della risposta anticorpale, che si attiva alcune settimane dopo l'infezione (1 settimana - 3 mesi), perciò tale intervallo è detto "periodo finestra", mentre la comparsa degli anticorpi è detta "sieroconversione".

In seguito ad un'efficace risposta immunitaria la viremia nel sangue si abbassa drasticamente; questo probabilmente non riflette un'inibizione della replicazione virale, ma piuttosto un'efficiente azione del sistema nel rimuovere le particelle virali che si sono prodotte. Durante la fase acuta si assiste inoltre ad un calo sistemico e drastico dei linfociti CD4+, che però è solo transitorio, poiché si verifica una normalizzazione dei valori una volta attivata la risposta anticorpale.

L'infezione acuta decorre spesso in modo asintomatico oppure può tradursi, entro 2-6 settimane dal contagio, in sintomi simili a quelli di una sindrome influenzale o mononucleosica: febbre, malessere, ingrossamento dei linfonodi, stanchezza, esantema orticarioide.

Latenza clinica

Lo sviluppo della risposta immunitaria specifica e la riduzione del carico virale nel sangue periferico segnano il passaggio dalla fase acuta a quella cronica, clinicamente latente, caratterizzata dalla mancanza di sintomi. A questa fase, che può durare diversi anni (dagli 8 ai 12), non corrisponde una latenza biologica, dal momento che la replicazione del virus continua negli organi linfoidi.

Possono perciò coesistere cellule in cui è presente solo il provirus integrato e cellule in cui invece sono presenti le diverse fasi della replicazione virale.

Il sangue periferico perciò non riflette precisamente l'effettivo stato immunologico dell'intero organismo, ma sono gli organi linfoidi (soprattutto linfonodi e cellule dendritiche follicolari) la sede principale della replicazione e propagazione virale. In questa fase si assiste inoltre a una progressiva diminuzione del numero di linfociti CD4+ circolanti e alla comparsa di difetti funzionali a carico sia dei linfociti CD4+ e CD8+, sia dei monociti e dei macrofagi, a cui si associa un progressivo deterioramento del sistema immunitario.

In assenza di trattamento farmacologico la maggioranza dei pazienti evolve verso la malattia (in genere quando il livello di linfociti CD4+ scende sotto il

valore 350/mm³). Una quota minore presenta un'evoluzione più rapida (rapid progressor) e una porzione, costituita da meno del 10% dei soggetti infetti, ha una tendenza a non ammalarsi anche dopo 12 anni di infezione²³.

Fase sintomatica o AIDS conclamato

Il progredire dell'immunodepressione si riflette, come detto, in una diminuzione dei linfociti CD4⁺ (nelle fasi finali meno di 200 linfociti/microlitro) e in difetti funzionali soprattutto in monociti-macrofagi, mentre la viremia (che in fase asintomatica si era assestata tra 1000 e 10000 copie RNA/ml) aumenta progressivamente fino a 10⁶, 10⁷ copie RNA/ml. Questo deterioramento conduce alla fase clinicamente sintomatica della malattia che si manifesta inizialmente con linfoadenopatia persistente (stadio "LAS", Lympho-Adenopatic Syndrome), seguito da una fase di perdita di peso, astenia, anemia, febbre, diarrea, ipergammaglobulinemia, che è definita stadio "ARC", ovvero AIDS-related complex. L'insorgenza di infezioni opportunistiche (fase di AIDS conclamato)²⁴ sono la principale causa di morte; tra queste le più frequenti sono:

toxoplasmosi viscerale, polmonite da *Pneumocystis carinii* (colpisce il 75% dei pazienti) e da Citomegalovirus, infezioni erpetica disseminata. Tutte queste infezioni sono causate da microrganismi ubiquitari dell'ambiente, normalmente poco patogeni in soggetti immunocompetenti, ma capaci di indurre manifestazioni patologiche gravi in condizioni di immunodepressione.

Caratteristiche dell'AIDS conclamato sono anche affezioni neoplastiche quali il Sarcoma di Kaposi, linfomi B cellulari, linfomi e carcinomi del collo dell'utero. Un organo in cui si assiste a una forte replicazione virale è il cervello, circa 2/3 dei pazienti affetti da AIDS presenta infatti lesioni degenerative a carico del Sistema Nervoso Centrale; il virus può infatti colpire le cellule del SNC quali macrofagi e cellule della microglia, portando a una forma di encefalopatia detta ADC (AIDS-dementia complex). In conclusione la patologia dell'infezione da HIV risulta caratterizzata da una irreversibile compromissione della risposta immunitaria dovuta principalmente al calo dei linfociti T helper, che si traduce nella comparsa incontrollata di infezioni opportunistiche, manifestazioni neoplastiche, lesioni degenerative del SNC e di alterazioni emopoietiche periferiche che determinano una grave compromissione del sistema emopoietico.

Tuttavia, benché i linfociti T⁴ siano il bersaglio principale dell'infezione, il numero di tali cellule realmente infette si mantiene sostanzialmente esiguo rispetto alla popolazione totale. Risulta perciò difficile comprendere l'estesa

distruzione di tali cellule, considerando il fatto che esse, in condizioni fisiologiche, vengono continuamente rimpiazzate dai progenitori ematopoietici della serie linfoidi. Per spiegare questa perdita di cellule è stata presa in considerazione l'apoptosi^{25,26}, un fenomeno fisiologico selettivo che in corso di AIDS viene esasperato.

Oltre ai meccanismi patologici basati su proteine in grado di determinare la morte cellulare programmata, il virus è in grado di innescare segnali alternativi: la proteina Tat infatti, liberata nell'ambiente dalle cellule infette, amplifica il danno cellulare tramite induzione, nelle cellule non infette, di una massiva produzione di citochine, che determina uno squilibrio dei sistemi omeostatici e di controllo della proliferazione e differenziazione cellulare.

La causa delle lesioni del SNC non è da ricercare nell'infezione dei neuroni, ma sembra sia una conseguenza della liberazione di gp120 da parte delle cellule gliali accessorie, che innesca l'apoptosi dei neuroni stessi. Vari altri fattori sono stati poi chiamati in causa come possibili potenziatori dell'immunodeficienza, come infezioni da virus (HIV-2, CMV, HSV-6 e 7) o da batteri^{24,25}.

Diversamente da quanto accade nei soggetti adulti, nei neonati infetti (che rappresentano circa il 10-40% dei nati da madri sieropositive) e nei bambini il virus tende ad essere più aggressivo e l'evoluzione dell'infezione molto più rapida, portando alla morte del paziente nell'arco di 5-6 anni. A questo quadro si

aggiunge poi la minor disponibilità di farmaci anti-HIV somministrabili a bambini di età inferiore a 13 anni. Per quanto riguarda il decorso clinico nei bambini il virus manifesta i suoi sintomi fin dal primo anno di vita, esso infatti infetta frequentemente il cervello ancora in maturazione, ostacolando lo sviluppo intellettuale, le funzioni motorie e causando problemi di coordinazione. Si possono avere inoltre disturbi nella crescita e una maggior suscettibilità alle infezioni batteriche, essendo il sistema immunitario del bambino non ancora pienamente formato e depresso, allo stesso tempo, dall'infezione virale. La maggior aggressività del virus nei bambini sembra essere in relazione alla costante ed elevata carica virale, che non subisce la diminuzione tipica dell'adulto; l'evoluzione della malattia risulta perciò più rapida e correlata a una ridotta sopravvivenza.

La diagnosi di infezione da HIV

Il virus HIV causa un'infezione permanente nel soggetto contagiato, perciò per effettuare una diagnosi di infezione è sufficiente rilevare la presenza

anticorpale specifica per il virus nel siero del paziente (sieropositività), senza dover ricercare l'agente patogeno. Lo schema generale del protocollo diagnostico prevede dapprima il test per lo screening di base, costituito dalla ricerca di anticorpi anti-HIV nel siero mediante saggio immunoenzimatico ELISA (figura 9), in cui gli antigeni virali (lisato virale, peptici ricombinanti gag ed env e/o peptici sintetici codificati dal gene env) sono adesi su fase solida. Attualmente sono disponibili saggi ELISA di quarta generazione che combinano la ricerca degli antigeni virali a quella degli anticorpi, riducendo così il "periodo finestra".

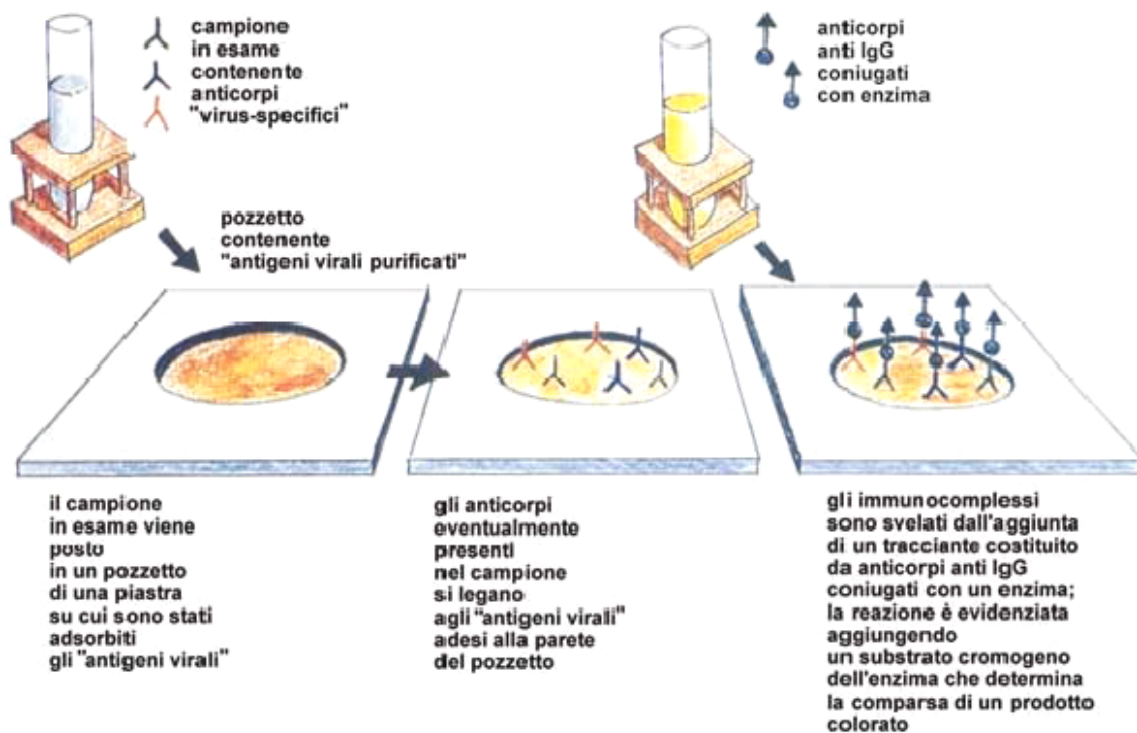


Figura 9. Test ELISA

Gli anticorpi specifici per HIV possono essere rilevati entro 2-8 settimane dopo l'infezione. Il tempo necessario per osservare la sierconversione dipende da diversi fattori, come la concentrazione infettante, la modalità di trasmissione e, infine, la sensibilità del test usato. La sierconversione inizia generalmente con la comparsa di IgM specifiche per la proteina Gag, poi, nell'arco di 5-7 giorni, compaiono le IgG dirette verso le proteine virali p24 e gp120, che però in questa fase sono difficilmente individuabili con i comuni saggi sierologici. Il titolo anticorpale delle IgG specifiche per le proteine di HIV aumenta progressivamente nei primi mesi, per poi stabilizzarsi raggiungendo un plateau.

Campioni positivi o indeterminati con il test immunoenzimatico ELISA vengono quindi sottoposti ad un test qualitativo di secondo livello, cioè Immunoblotting o Western blotting (**figura10**), che consente di evidenziare la presenza di anticorpi contro le singole proteine virus-specifiche, separate in base al peso molecolare.

Per essere considerato positivo a questo test un campione di siero deve possedere anticorpi verso almeno una proteina del core (p24) ed una dell'envelope (gp41, gp120 o gp160).

Gli anticorpi specifici per HIV possono essere rilevati entro 2-8 settimane dopo l'infezione. Il tempo necessario per osservare la sierconversione dipende da diversi fattori, come la concentrazione infettante, la modalità di trasmissione e,

infine, la sensibilità del test usato. La sieroconversione inizia generalmente con la comparsa di IgM specifiche per la proteina Gag, poi, nell'arco di 5-7 giorni, compaiono le IgG dirette verso le proteine virali p24 e gp120, che però in questa fase sono difficilmente individuabili con i comuni saggi sierologici. Il titolo anticorpale delle IgG specifiche per le proteine di HIV aumenta progressivamente nei primi mesi, per poi stabilizzarsi raggiungendo un plateau. Campioni positivi o indeterminati con il test immunoenzimatico ELISA vengono quindi sottoposti ad un test qualitativo di secondo livello, cioè Immunoblotting o Western blotting (**figura10**), che consente di evidenziare la presenza di anticorpi contro le singole proteine virus-specifiche, separate in base al peso molecolare. Per essere considerato positivo a questo test un campione di siero deve possedere anticorpi verso almeno una proteina del core (p24) ed una dell'envelope (gp41, gp120 o gp160).

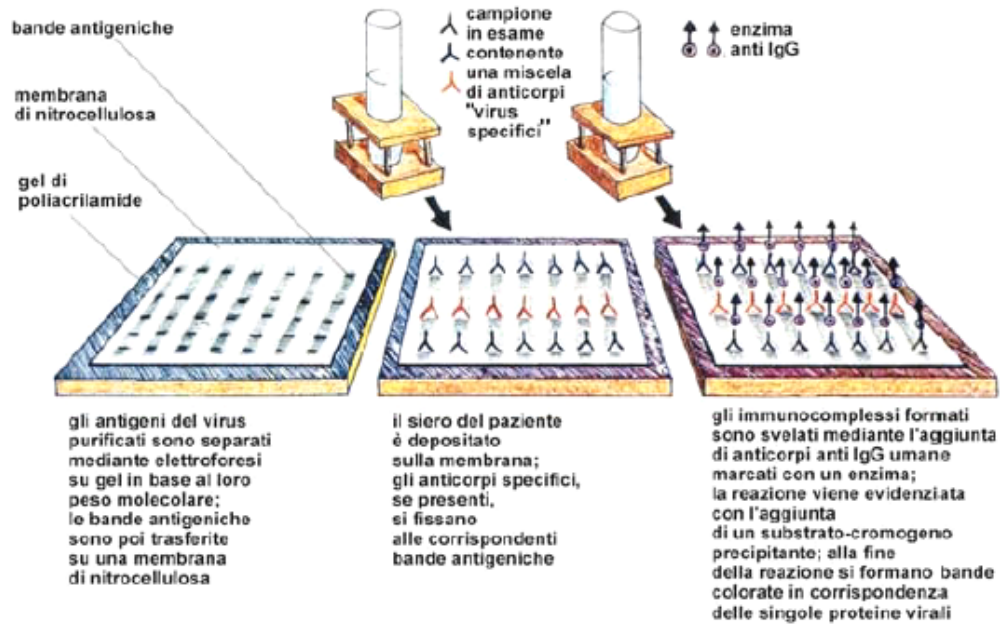


Figura 10. Western Blot

Nei casi di reattività verso una sola proteina virale si ha un risultato indeterminato e perciò serve un secondo controllo dopo 15-30 giorni.

Pur consentendo, di norma, una sicura diagnosi di infezione, la ricerca di anticorpi ha limiti di utilizzo in alcune circostanze:

- a) nella fase iniziale dell'infezione ("periodo finestra" di 3-4 settimane), poiché la quantità degli anticorpi circolanti non è ancora sufficiente ad essere evidenziata con le tecniche sierologiche;

- b) nei bimbi nati da madri infette, poiché possiedono anticorpi sierici anti-HIV di origine materna (IgG), di nessun valore diagnostico;
- c) nei soggetti che dalle indagini sierologiche risultano “borderline”, ovvero i cui risultati sono di dubbia positività.
- d) Nei casi in cui non si può ricorrere ad indagini sierologiche o in caso di risultati indeterminati di queste ultime, è necessario ricorrere ad indagini virologiche in grado di evidenziare la presenza del virus in forma attiva o latente.

Questa ricerca può essere effettuata attraverso l'isolamento del virus in colture cellulari in vitro (ma questa tecnica è indaginosa e richiede tempi piuttosto lunghi e laboratori attrezzati), attraverso la dimostrazione della presenza di antigeni virus-specifici, soprattutto della proteina p24 del core virale, o di specifiche sequenze nucleotidiche, in particolare DNA provirale nei linfomonociti circolanti e l'RNA virionico nel plasma. La ricerca di p24 e dell'RNA virale nel plasma assicurano un'elevata sensibilità nella diagnosi, ma per essere certi di un'infezione in atto la ricerca del DNA provirale nei linfomonociti circolanti, attraverso metodiche di amplificazione quali la PCR, è in assoluto la metodica che dà maggior affidabilità, rilevando la presenza di virus

latente nella cellula (provirus) indipendentemente dalla replicazione virale e permettendo di ottenere una diagnosi in tempi più rapidi, senza dover manipolare virus infettante^{29,30}.

Questa metodica risulta utile anche nei casi di bambini nati da madre sieropositiva, nei quali per esempio è indispensabile distinguere tra gli anticorpi autoctoni e IgG di origine materna, che possono persistere per 6-24 mesi e che perciò rendono impossibile l'uso dei test sierologici in bimbi di età inferiore ai 18 mesi. Per questi soggetti vengono perciò eseguiti test virologici utilizzando la PCR.

Nel primo mese di vita possono esserci dei falsi positivi, ma comunque già alla nascita un 38% dei bambini infetti risultano positivi alla PCR e la sensibilità della metodica aumenta, fino a raggiungere il 100% in 14-30 giorni²⁷⁻²⁸.

La trasmissione del virus

La trasmissione del virus HIV avviene attraverso il contatto con il sangue di un soggetto infetto e per via sessuale. Per tale motivo, l'AIDS viene annoverata nel gruppo delle malattie a trasmissione sessuale (MST). L'infezione si diffonde più rapidamente tra individui che hanno spesso rapporti sessuali non protetti con partner diversi e tra i tossicodipendenti. Il virus può anche essere trasmesso dalla madre sieropositiva al feto, attraverso la circolazione sanguigna placentare, o al bambino dopo la nascita, attraverso:

- allattamento al seno
- contagio per via sessuale
- contagio per via sessuale

Nel caso del contagio per via sessuale, l'HIV presente nello sperma e nelle secrezioni vaginali si immette nella circolazione sanguigna del partner non infetto attraverso piccole abrasioni delle mucose (genitali o orali), già presenti o

formatesi durante i rapporti sessuali. Il contagio si verifica allo stesso modo negli individui omosessuali ed eterosessuali.

- Contagio tra tossicodipendenti

La trasmissione del virus tra tossicodipendenti riguarda coloro che fanno uso di droghe iniettabili, come l'eroina, e che impiegano siringhe già usate; in tal caso, anche piccole quantità di sangue depositatesi sull'ago o aspirate al momento dell'estrazione della siringa, possono essere sufficienti ad infettare il successivo utilizzatore di questa.

- Altre vie di trasmissione

Particolari vie di contagio sono quelle che si stabiliscono tra pazienti portatori di HIV e operatori sanitari, e viceversa, e nel corso di trasfusioni sanguigne. La probabilità di contrarre l'infezione per queste vie in realtà è piuttosto bassa, per le misure di prevenzione e le condizioni di sterilità adottate in ambito sanitario e grazie ai test di routine per l'individuazione dell'HIV

effettuati nelle emoteche. La donazione del sangue non comporta per il donatore alcun rischio.

Non vi è prova che l'HIV possa essere trasmesso attraverso l'aria, le punture di insetti, il sudore, la saliva, oppure tramite il contatto con persone infette, purché non vi sia scambio di sangue o di secrezioni genitali: dunque, il virus non si diffonde con una stretta di mano, o impiegando gli stessi attrezzi da lavoro di un sieropositivo, o indossandone un abito. Ciò è dovuto al fatto che l'HIV non sopravvive a lungo quando viene esposto all'ambiente. Invece, la condivisione di oggetti come rasoi, spazzolini da denti, bende, non è immune dal rischio di contagio.

Sintomatologia dell'AIDS

I primi sintomi dell'AIDS sono simili a quelli che si sviluppano in soggetti con un normale sistema immunitario. La maggior parte sono infezioni causate da batteri, virus, funghi, parassiti e altri organismi. Negli individui affetti da AIDS sono comuni le infezioni opportunistiche, e aumenta il rischio di sviluppare varie

forme di tumore come il Sarcoma di Kaposi, tumori del cervello e linfomi. Sintomi comuni sono febbre, sudorazione specie notturna, ingrossamento ghiandolare, tremore, debolezza e perdita di peso²⁸. La sopravvivenza media con terapia antiretrovirale è di 4-5 anni dal momento della diagnosi di AIDS conclamato . Senza il supporto terapeutico la morte sopravviene entro un anno.²⁹ La maggior parte dei pazienti muore per infezioni opportunistiche dovute al progressivo indebolimento del sistema immunitario.

La terapia farmacologica

Fino a qualche tempo fa l'intervento medico nella terapia dell'infezione da HIV-1 si limitava al controllo e alla cura delle infezioni opportunistiche che insorgono come conseguenza della malattia. Negli ultimi anni sono stati fatti enormi passi avanti e questo tipo di trattamento è stato completamente sostituito, grazie all'introduzione di farmaci antiretrovirali.

La miglior comprensione della patogenesi dei danni prodotti dal virus, la possibilità di determinare la carica virale e di avere perciò un parametro diretto

dell'effettiva replicazione virale e la disponibilità di nuovi farmaci a potente attività antiretrovirale, ha spostato l'attenzione medica da un ambito di tipo preventivo verso la ricerca di combinazioni terapeutiche a scopo "curativo", il più possibile efficaci e al tempo stesso tollerabili dal paziente. Oggi perciò lo scopo primario della terapia è abbassare la carica virale quando il sistema immunitario è ancora efficiente e quindi in grado di recuperare pienamente le sue funzioni. Poiché il virus dipende per la sua replicazione dal macchinario biologico della cellula ospite, i farmaci antiretrovirali devono inibire selettivamente la replicazione virale senza danneggiare il metabolismo cellulare.

Gli studi di virologia molecolare hanno permesso di identificare funzioni virali specifiche:

- l'attacco del virus alla cellula target
- la trascrizione inversa del genoma virale
- la traduzione delle proteine e successivo assemblaggio con la maturazione della progenie virale che possono servire da veri bersagli per l'inibizione.

Tuttavia risulta difficile sviluppare agenti antivirali in grado di distinguere i processi replicativi del virus da quelli della cellula ospite.

Attualmente sono registrati in Italia farmaci antiretrovirali appartenenti a tre classi farmacologiche, ognuna con un diverso meccanismo d'azione³⁰⁻³¹.

Tutti questi farmaci non sono in grado di uccidere il virus, ma solo di bloccarne la replicazione, perciò i pazienti in trattamento, anche se hanno una carica virale non rilevabile nel sangue, devono comunque considerarsi sempre potenzialmente infettanti. Qualsiasi decisione sull'inizio o il cambiamento di terapia dovrebbe essere guidata dal continuo monitoraggio di parametri quali la carica virale e il numero di CD4⁺.

Analoghi Nucleosidici della Trascrittasi Inversa (NRTI)

I più noti sono la zidovudina (AZT), la didanosina (DDI), la zalcitabina (DDC), la lamivudina (3TC), la stavudina (D4T) e l'abacavir. Per i primi dieci anni dalla scoperta del virus, questi farmaci sono stati i più usati in terapia. Sono detti nucleosidici per la loro somiglianza strutturale con i nucleosidi trifosfati. Agiscono nelle fasi precoci della replicazione per prevenire l'infezione delle cellule sane; dopo il loro ingresso nella cellula questi farmaci vengono fosforilati dalle chinasi cellulari e sono così convertiti nella forma trifosfato attiva e competono con il deossinucleotide naturale per il legame all'RT. Competono inoltre anche

coi dideossinucleotide trifosfato per l'incorporazione nella nuova catena nascente di DNA, bloccandone in tal modo la sintesi ulteriore (sono terminatori di catena poiché manca loro l'ossidrile in posizione 3' o alterandone la funzione. Sicuramente il farmaco più studiato è la zidovudina, utilizzato sin dal 1987. È l'unico in grado di passare la barriera emato-encefalica e perciò di prevenire le manifestazioni neurologiche dell'ADC (AIDS-Dementia Complex) e di prevenire il passaggio da madre a figlio durante la gravidanza.

Analoghi non-Nucleosidici della Trascrittasi Inversa (NNRTI)

Questa classe di farmaci fu scoperta circa 10 anni fa, ma il loro sviluppo è stato ostacolato dagli scarsi risultati ottenuti in seguito all'impiego in monoterapia, per la rapida insorgenza di resistenza. I più noti sono la nevirapina, la delaviridina, l'efavirenz. Questi farmaci sono inibitori non competitivi, altamente selettivi per la trascrittasi inversa. Per essere attivati non richiedono la fosforilazione e non competono con i nucleosidi trifosfato, dei quali infatti non imitano la struttura. Il meccanismo con cui riescono a bloccare la trascrittasi è di tipo allosterico: agiscono legandosi direttamente al sito attivo dell'enzima, bloccandone l'azione ed impedendo così che avvenga la formazione di DNA

provirale. Hanno il vantaggio di essere privi di effetti sulle cellule ematopoietiche e di non mostrare resistenza crociata con gli inibitori nucleosidici dell'RT.

Inibitori della Proteasi (PI)

A differenza delle classi precedenti che lavorano in una fase precoce del ciclo replicativo, questi farmaci inibiscono la proteasi virale, che è indispensabile in uno stadio tardivo del ciclo per scindere i precursori proteici in proteine virali mature.

Questi farmaci sono composti a basso peso molecolare che interagiscono stericamente col sito attivo dell'enzima, situato all'interfaccia delle due subunità identiche che lo compongono e sono tutti composti altamente idrofobici, proprio per poter interagire col sito catalitico della proteasi, inibendone reversibilmente l'attività.

Poiché l'azione enzimatica della proteasi si esplica attraverso il processamento dei precursori della retrotrascrittasi, della integrasi e della proteina Gag e che solo le forme processate di queste proteine sono utilizzate per la costruzione di particelle virali infettanti, si deduce che questi farmaci determinano la produzione di particelle virioni che difettive, incapaci di infettare

nuove cellule. Fanno parte degli inibitori della proteasi il saquinavir, il ritonavir, l'indinavir, il nelfinavir e l'amprenavir.

Recentemente si è pensato di intervenire impedendo l'ingresso del retrovirus nella cellula, andando ad agire nella fase in cui HIV si lega alla cellula che andrà ad infettare³⁶.

È stata dunque introdotta una nuova categoria di farmaci, gli **inibitori della fusione**, a cui appartiene T-20 (Enfuvirtide), un peptide di sintesi derivato dalla proteina virale transmembrana gp41. Esso è in grado di legarsi alla glicoproteina gp41, impedendone il legame del virus con la cellula bersaglio e quindi bloccandone l'ingresso.

Il fallimento terapeutico

Ogni farmaco porta alla selezione di mutanti virali resistenti e ciò è la causa principale dell'insuccesso terapeutico e l'ostacolo maggiore all'impiego del composto per trattamenti protratti di un'infezione virale a lungo decorso. Poiché

le mutazioni compaiono durante la replicazione, la loro insorgenza è la diretta conseguenza di un'incompleta soppressione dell'attività virale che si verifica o per una scorretta assunzione del farmaco o per inadeguatezza della terapia stessa. Per ridurre il rischio dell'insorgere di resistenze è necessario ottimizzare l'aderenza del paziente alla terapia e utilizzare più farmaci in combinazione tra loro, ma è comunque documentato un fallimento terapeutico, nel primo anno di cura, nel 50% dei pazienti che iniziano la terapia antiretrovirale³².

L'uso di cocktail di farmaci, di solito due analoghi dei nucleosidici e non nucleosidici e un inibitore della proteasi, oltre a rallentare lo sviluppo di farmacoresistenza, sembra in grado di ridurre la tossicità cronica del farmaco, poiché permette di ottenere lo stesso effetto con dosaggi inferiori dello stesso. Per ritardare l'insorgenza di mutazioni ci si basa sul fatto che in presenza di due farmaci la frequenza mutazionale per ciascuno non varia, ma diminuisce in modo esponenziale la possibilità di selezionare varianti con mutazioni resistenti per entrambi.

Se un farmaco non è abbastanza potente o se è assunto in concentrazione troppo bassa, il ceppo selvaggio del virus resta prevalente nell'organismo e la carica virale elevata, senza comparsa di ceppi resistenti. Molto importante è anche l'aderenza del paziente alla terapia; attualmente si ricorre a diverse assunzioni giornaliere con un preciso schema di somministrazione. I farmaci

hanno numerosi effetti collaterali quali eritema, cefalea, nausea, anemia, insensibilità agli arti inferiori etc., che spesso portano a un interruzione del trattamento, indipendentemente dai benefici clinici ottenuti. Una terapia è tanto più efficace quanto più essa è in grado di agire sui reservoir anatomici, dove vi è molte volte una diversa percentuale di mutazioni, e sui reservoir cellulari costituiti dai macrofagi, in cui il virus si replica persistentemente ed è protetto da fattori immunologici e farmacologici.

La Farmacoresistenza

La resistenza farmacologica si instaura a seguito di mutazioni nel genoma del virus, in particolar modo a livello del gene pol, bersaglio preferenziale dei più noti farmaci antiretrovirali. Tali mutazioni, in genere puntiformi, sono particolarmente frequenti durante la replicazione del virus, mentre alcune però sono compatibili con l'infettività del virus stesso (mutazioni silenti in terza base), altre, che intervengono in punti critici quali i geni per proteine strutturali,

regolatrici o enzimatiche, fanno sì che vengano così codificate proteine che svolgono ancora la loro funzione nel ciclo replicativo del virus, ma con un'affinità molto minore per il farmaco. Il numero di mutazioni necessarie per conferire farmacoresistenza è variabile e, in alcuni casi, il tipo e la sequenza di mutazione possono essere più importanti del numero stesso.

La frequente variabilità di HIV-1 costituisce perciò il principale fattore per l'emergere di varianti resistenti, in risposta alla pressione selettiva farmacologica. Durante la replicazione del genoma virale intervengono infatti tre sistemi enzimatici: trascrittasi inversa (virale), DNA polimerasi ed RNA polimerasi (cellulari), che compiono errori di copiatura tali da permettere al virus di superare gli ostacoli posti dalla pressione selettiva. Inoltre la trascrizione inversa non possiede l'attività di "proof-reading" e "correzione di bozze" ed è perciò più frequente l'incorporazione di basi errate, aumentando così la facilità nell'instaurarsi di mutazioni. Si calcola che, in vitro, la frequenza di errore dell'RT è di una base su 10⁴ sintetizzate³⁹.

Le mutazioni in punti critici come quelli sopra citati possono perciò incidere in modo drammatico sulla virulenza, la capacità di replicazione, la citotossicità e la risposta alla terapia antiretrovirale³³. Durante le prime fasi dell'infezione, la popolazione virale sembra essere monoclonale, con sequenze nucleotidiche relativamente omogenee. In seguito, durante il corso della malattia,

essa appare invece disomogenea, fino a raggiungere un milione di varianti nel soggetto asintomatico e più di 100 milioni di varianti nel paziente in fase avanzata. La popolazione virale è perciò costituita da una specie dominante (wild type) con la miglior fitness riproduttiva, e molte piccole popolazioni distinte geneticamente e deficitarie dal punto di vista replicativo, che rimangono di proporzioni ridotte fino a che non viene attivata una pressione selettiva operata dai farmaci.

La terapia farmacologica crea perciò una situazione dinamica con continua produzione di varianti che diventeranno nel tempo la specie dominante con conseguente selezione di ceppi resistenti ai farmaci utilizzati³⁴.

Un virus così mutato può infettare le cellule T attivate ed essere replicazione competente, oppure può integrarsi come provirus e restare latente per lungo tempo. La mancata rilevazione di mutazioni nell'RNA plasmatico quindi non è sufficiente per escludere la presenza di esse nel provirus.

La farmacoresistenza viene dunque definita come:

- a) la perdita della sensibilità del virus all'effetto del farmaco;
- b) la comparsa di un fenotipo peculiare di HIV;
- c) il risultato di un'alterazione del genotipo;
- d) la causa primaria del fallimento terapeutico.

È dunque necessario un controllo periodico della sensibilità/resistenza del soggetto in terapia verso le diverse classi di farmaci disponibili, per instaurare un adeguato regime terapeutico³⁵⁻³⁷. La farmaco-resistenza è causata da un fenotipo virale che rappresenta il risultato di un genotipo virale alterato, perciò l'analisi della sensibilità/resistenza nei confronti delle diverse classi di farmaci è possibile sia con metodi fenotipici, sia con metodi genotipici. Determinazione fenotipica della resistenza: consiste nel determinare la capacità del virus isolato dal paziente di moltiplicarsi in colture di cellule in vitro, mantenute con dosi adeguate del farmaco da saggiare.

Questi test sono interpretabili facilmente e con sicurezza, infatti se c'è moltiplicazione del virus nella coltura in presenza di farmaco si deduce che esso sia resistente (o almeno una quota apprezzabile della popolazione virale lo sia), mentre se non si osserva moltiplicazione è chiaro che la popolazione virale infettante è omogeneamente sensibile.

Questo tipo di test però richiede laboratori attrezzati per isolare e coltivare HIV in colture di cellule in vitro, l'esecuzione è complessa, i tempi sono lunghi e i costi elevati. Determinazione genotipica della resistenza: consiste nella ricerca diretta nel genoma virale, amplificato dal materiale patologico, di mutazioni correlate a farmaco-resistenza (ad oggi ne sono note almeno 100), identificate per confronto con una sequenza virale nota wild-type.

Questo tipo di test prescinde dall'impiego di colture cellulari, è relativamente poco costoso e i tempi necessari per ottenere una risposta utile molto più brevi (3-5 giorni) di quelli dei test fenotipici.

Per entrambi i tipi di test i livelli di sensibilità non sono estremamente elevati, infatti ambedue evidenziano la presenza di subset di virus portatori di mutazioni associate a farmaco-resistenza solo se presenti in almeno il 20% della popolazione virale in osservazione; inoltre sono efficienti partendo da una carica virale di almeno 1000 copie/ml di RNA³⁸.

Nonostante queste limitazioni lo studio delle farmaco-resistenze della popolazione virale presente nel paziente è uno degli esami essenziali da effettuare prima di iniziare o di modificare un trattamento farmacologico, al fine di limitare al massimo i possibili rischi di insuccesso terapeutico^{39,40}.

Parte II

Descrizione del target: l'enzima integrasi di HIV-1

Struttura e domini funzionali

Come già detto in precedenza, i retrovirus codificano i loro tre enzimi (proteasi, trascrittasi inversa ed integrasi) all'interno del gene *POL*, che è traslato come poliproteina Pol. La proteasi è codificata all'estremità 5' e l'integrasi all'estremità 3' del gene *POL*. L'integrasi è generata durante la maturazione del virus dalla scissione della poliproteina Pol ad opera dell'HIV proteasi. L'integrasi dell'HIV è una proteina del peso di 32 kDa e comprende tre domini strutturali: il dominio ammino terminale (NTD), il core catalitico (CCD) ed il dominio carbossi terminale (CTD) (**figura 11**). La struttura atomica di questi tre domini è stata determinata attraverso la diffrazione a raggi X o la risonanza magnetica nucleare (NMR). La **figura 11** mostra le strutture di CCD⁴³ in associazione con CTD⁴² e di NTD⁴¹.

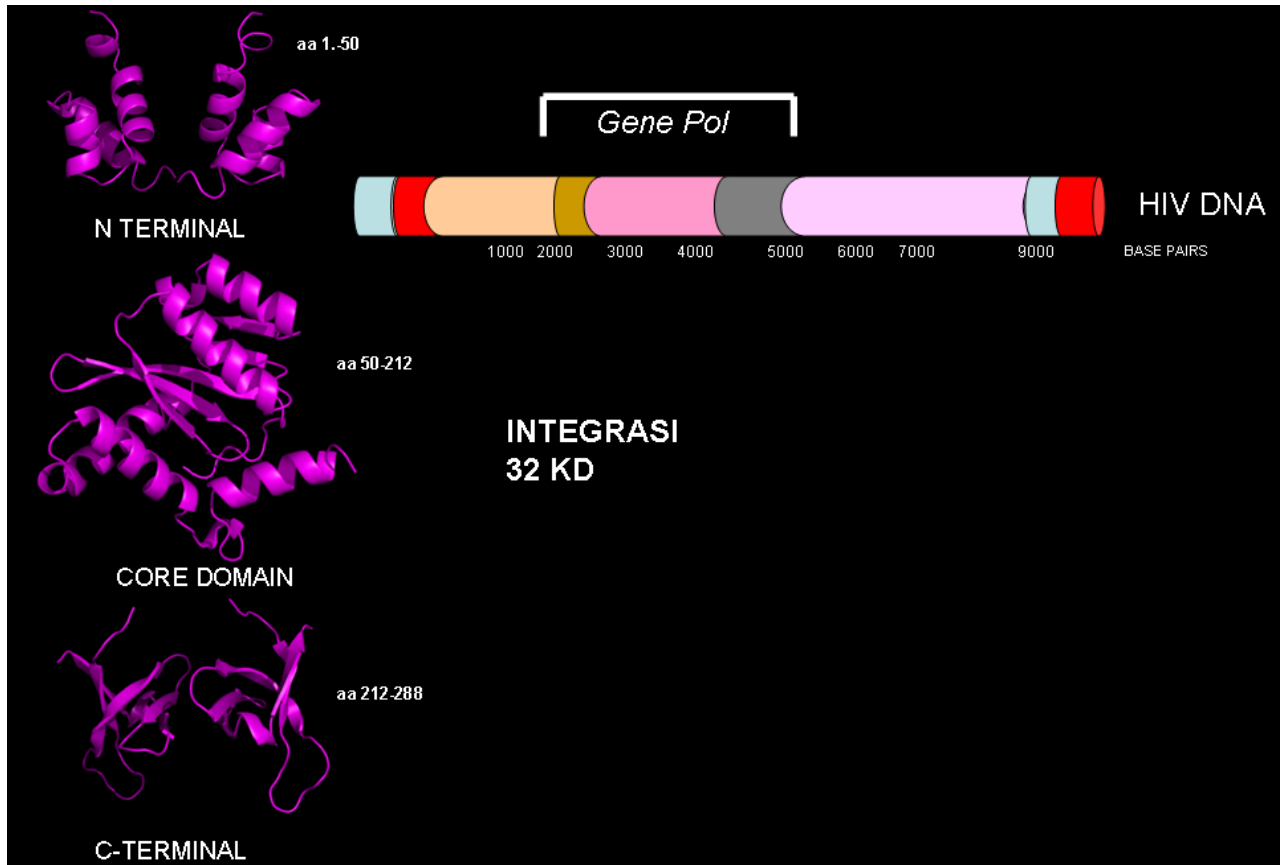


Figura 11. Struttura cristallina dell'N terminal domain (pdb code1E0E) del C- terminal domain (pdb code 1QMC) e del core domain (1-bl3) di IN

Il CCD, che comprende i residui 50-212 organizzati in cinque β -sheet intrecciati a regioni di α -eliche⁴⁴ (**figura 12**), forma un dimero in tutte le strutture esaminate, ed è strutturalmente simile a quello delle altre integrasi retrovirali (Murine Leukemia Virus (MLV) e Avian Sarcoma Virus (ASV)⁴⁵, come riportato in **figura 12**.



Figura 12. Rappresentazione del dominio catalitico in forma dimerica (codice PDB: 1ITG).

Questa famiglia di enzimi che processano il DNA (polinucleotidil transferasi) contiene una tripletta di amminoacidi caratteristica (motivo DDE), formata nell'integrasi dell'HIV dalla triade catalitica D64 (aspartato 64), D116 (aspartato 116) e E152 (glutammato 152) (**figura 13**). Questi residui sono altamente conservati in tutte le integrasi e le retrotransposasi (motivo DDE). La mutazione, anche di uno solo di questi tre amminoacidi, annulla l'attività enzimatica e la replicazione virale. I due residui D64 e D116 formano un

complesso di coordinazione (con legami covalenti) con uno ione bivalente (Mn^{2+} o Mg^{2+}).



Figura 13. Rappresentazione del dominio catalitico del CCD di IN (codice PDB: 1bl3 chain A).

Poichè nella struttura cristallina dell'integrasi di ASV è stato osservato un altro metallo bivalente⁴⁶, e poichè nella struttura delle altre polinucleotidil transferasi sono presenti due cationi^{47,48} è stato proposto che possa essere coordinato un secondo metallo (Mn^{2+} o Mg^{2+}) a livello dei residui D116 e E152, quando l'integrasi lega il suo DNA substrato^{49,50}. È quindi probabile che i metalli coordinino sia l'integrasi che lo scheletro fosfodiesterico del DNA substrato durante

il 3'-processing e lo strand transfer. In più strutture, il CCD contiene un piccolo loop disordinato (residui 141-150), la cui struttura potrebbe essere stabilizzata dal DNA. Sebbene il CCD contenga il sito catalitico, in assenza dei domini NTD e CTD, esso riesce a catalizzare *in vitro* esclusivamente la reazione di disintegrazione, cioè l'inverso della reazione di strand transfer^{51,52}. Comunque non è ancora nota l'importanza fisiologica di tale reazione. La disintegrazione resta la sola reazione catalizzata dal CCD isolato: per catalizzare il 3'-processing e lo strand transfer, infatti, esso ha bisogno sia del dominio NTD che di quello CTD uniti a formare un complesso dimerico⁵³. Il dominio NTD consiste in un ammasso di tre α -eliche, comprende i residui 1-50 e contiene il motivo HHCC (**figura 2b**), comune a tutte le integrasi retrovirali. Il legame di un atomo di Zn^{2+} sembra indispensabile alla stabilizzazione della struttura del dominio NTD ed è necessario all'attività integrasica. Singole mutazioni a carico di questi quattro amminoacidi riducono l'attività enzimatica dell'integrasi⁵⁴. L'interfaccia del dimero del NTD è diversa nella struttura cristallina (**figura 4**) rispetto alla struttura determinata con l'NMR in soluzione: ciò indica un arrangiamento variabile dei multimeri dell'integrasi. Il dominio CTD, che comprende i residui 212-288, ha un folding di tipo β -strand SH3 e lega il DNA in modo non specifico⁵⁵, ma è indispensabile per le attività di 3'-processing e strand transfer. La **figura 4** mostra le strutture dei domini NTD e CTD. Queste strutture

indicano la possibilità che il dominio NTD sia posizionato tra il CCD e il CTD, vicino all' α -elica che lega il CCD al CTD. L'analisi della struttura dell'interfaccia tra i due domini all'NMR, che è differente da quella osservata all'analisi cristallografica, potrebbe far supporre che l'integrasi si organizzi in complessi multimolecolari (tetrameri e/o ottameri), come sembra accadere per l'enzima nella sua forma attiva. Sebbene la struttura di ciascun dominio sia stata determinata, la struttura tridimensionale dell'integrasi "full-length" non è tuttora disponibile. Il relativo arrangiamento spaziale dei domini strutturali dell'integrasi e le loro interazioni con il DNA substrato sono rimaste quasi completamente sconosciute: lo studio diretto viene infatti impedito dalla tendenza a formare grandi aggregati molecolari nelle condizioni di reazione. Sono state quindi determinate le strutture cristalline che comprendono il dominio catalitico assemblato o con il dominio N-terminale (IN1-212) o con il dominio C-terminale (IN52-288). Queste strutture sperimentali, coerenti con le strutture di ciascun dominio, determinate tramite cristallografia a raggi X, forniscono maggiori informazioni sulla disposizione spaziale nel contesto di una struttura statica e completa dell'integrasi di HIV-1 in assenza del DNA substrato. Questi studi dimostrano che un oligomero, probabilmente un tetramero dell'integrasi, è indispensabile per una completa reazione di integrazione. Due dei quattro siti attivi sono coinvolti nella reazione di integrazione, mentre gli altri due siti attivi

fungono da supporto strutturale per il tetramero. In particolare la struttura del *core* catalitico e del dominio N-terminale (IN1-212) formano un dimero di dimeri, che rappresenta un modello dell'arrangiamento tetramerico dell'integrasi (**figura 13**). In questa disposizione tetramericata, i due siti attivi adiacenti, uno per ciascun dimero, tendono ad avvicinarsi attraverso modificazioni conformazionali all'interno di ciascuna unità dimerica, allo scopo di facilitare le reazioni di integrazione. Mentre le informazioni statiche strutturali sono difficilmente valutabili, è stato invece osservato, attraverso simulazioni molecolari dinamiche, che la flessibilità di questo "loop" catalitico ha una notevole incidenza sulla funzione e sull'attività catalitica dell'HIV-1 integrasi⁵⁶.

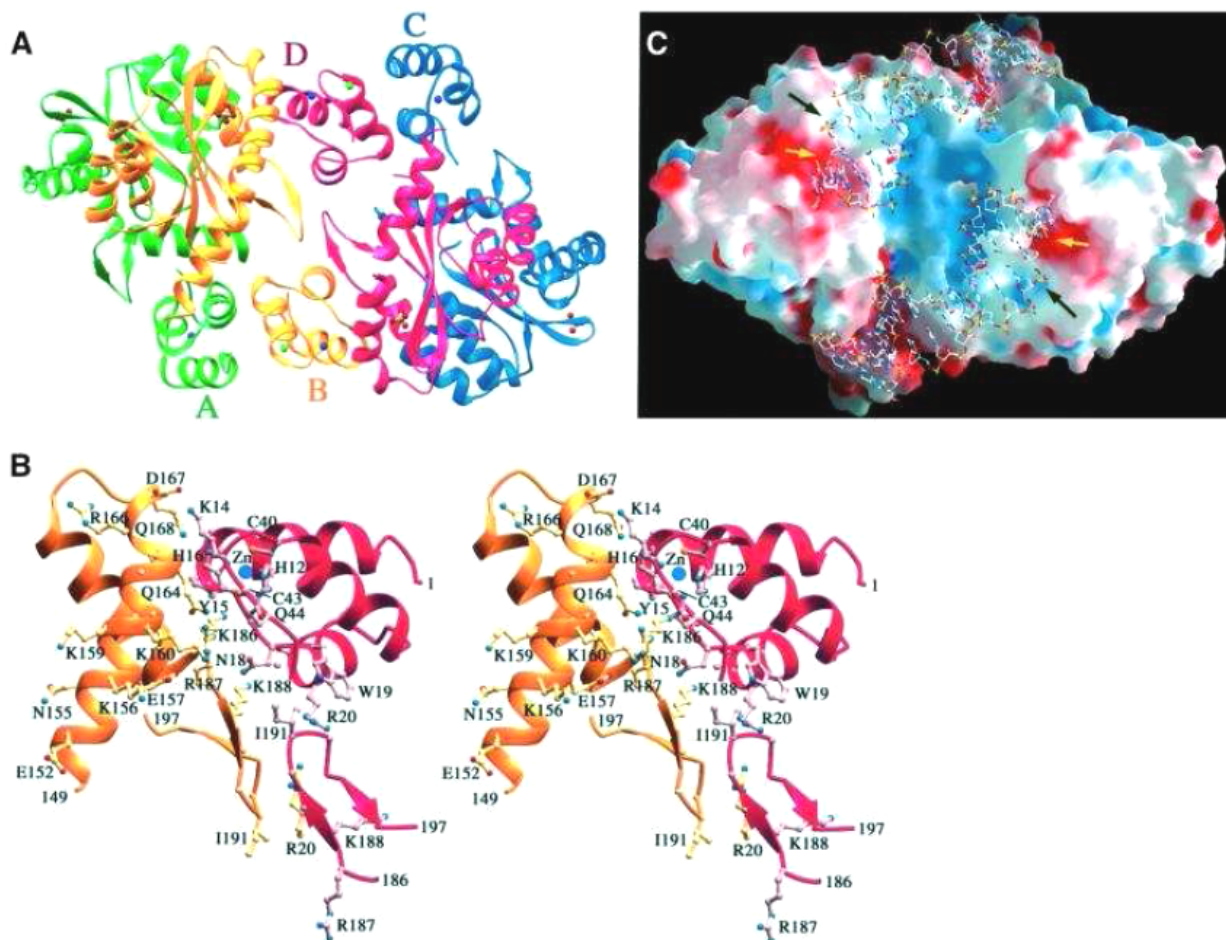


Figura 13. Il tetramero ABCD di IN1-212

I siti attivi delle subunità B e D del tetramero ABCD giacciono su una superficie concava con un potenziale elettrostatico che sembra essere favorevole all'interazione con il DNA (figura 13C) e ricordano i due siti attivi trovati nella Tn5 transposasi .

L'altra coppia di siti attivi giace sulla superficie opposta convessa del tetramero (distanza $> 90\text{\AA}$) e le subunità che hanno questi siti attivi probabilmente giocano un ruolo di supporto strutturale. I siti attivi B e D del tetramero dell'IN1-212 dell'HIV-1 sono separati da circa 40\AA , la stessa della Tn5 transposasi. Questa distanza è ancora maggiore dello spazio di cinque basi azotate dei siti d'intersezione delle due estremità del DNA virale e del DNA bersaglio, ma i due siti attivi sono plausibilmente tenuti più vicini dalle interazioni dell'interfaccia del dimero. Comunque i siti attivi si trovano sulla stessa superficie concava del complesso, semplificando molto le variazioni conformazionali necessarie a tenerli abbastanza vicini per la reazione finale di strand transfer. In alternativa l'inserimento può verificarsi in sequenza, assistito da un riorientamento proteico. Il tetramero ABCD, sebbene non osservato in soluzione con IN1-212 HIV-1, può esistere con la proteina "full-length". Il doppio dominio proteico C terminale (IN50- 288) forma sia i dimeri che i tetrameri in soluzione⁵⁷ facendo pensare ad un ruolo cruciale per il dominio C terminale nella tetramerizzazione. Il dimero dell'integrasi HIV-1 "full-length", risultante dalla sovrapposizione delle due strutture dei domini N terminale e C terminale, può accomodarsi in questo dimero di dimeri senza ingombri sterici nella catena principale. Per di più i domini C terminali delle subunità A e C si estendono sulla superficie convessa formando una spirale avvolta tra le eliche e

stabilizzano potenzialmente il tetramero. La risultante struttura cuoriforme tetramerica dell'integrasi assomiglia marcatamente alla struttura della Tn5 transposasi. L'altra coppia di domini C-terminali dalle subunità B e D si proiettano dalla superficie legante il substrato definita dai siti attivi estesi ed ampliano la zona caricata positivamente all'interfaccia dimero dimero (**figura 7C**) e quindi, potenzialmente, partecipano al legame dei terminali virali. Il tetramero ABCD formato senza il dominio C terminale è stato osservato soltanto in cristalli; si ipotizza che un'integrasi tetramerica attiva possa essere simile in tutto a questa configurazione tranne degli aggiustamenti all'interfaccia del tetramero.

Modelli di interazione dell'integrasi dell'HIV-1 con il DNA

La natura non specifica del legame con il DNA, sia virale che dell'ospite, rappresenta una sfida ancora maggiore per giungere alla determinazione della struttura di un complesso. Comunque, la struttura cristallina di IN1-212, insieme con altri dati sia biochimici che strutturali, pone dei vincoli sui modelli praticabili. La presenza di un fosfato posto a circa 7Å dai residui del sito attivo,

che eguaglia la distanza tra i fosfati adiacenti lungo il backbone del DNA, in tutti e quattro i monomeri nell'unità asimmetrica, identifica un potenziale sito di legame del fosfato vicino ai terminali del DNA virale. Invece questo fosfato interagisce con la catena laterale dell'amminoacido K159, un residuo che è stato identificato mediante photo-cross-linking perché è vicino all'adenina del dinucleotide CA adiacente al fosfato⁵⁸.

Nella struttura della Tn5 Transposasi complessata con il substrato⁵⁹, il fosfato in 5' è in contatto con K333, un equivalente di K159 dell'HIV-1 integrasi. Perciò è possibile ipotizzare che il fosfato posto in questa struttura mimica il fosfato più vicino al 5' del legame fosfodiesterico. L'orientamento del DNA virale nel sito attivo è anche vincolato dalla localizzazione dei residui del sito attivo e dalla probabile interazione dei due nucleotidi terminali all'estremità 5' del DNA virale con i residui in vicinanza di Q148. L'incubazione dell'integrasi dell'HIV-1 con il substrato virale, in presenza di uno ione metallico divalente, termina nella formazione di un complesso stabile che è resistente alla competizione con il DNA concorrente⁶⁰. La rimozione dei due nucleotidi in 5' dal substrato virale⁶⁰ o la mutazione di Q148 a leucina⁶¹ blocca la formazione del complesso, suggerendo che le due entità interagiscono normalmente. Posti questi vincoli, l'orientamento del substrato virale nella regione del sito attivo può essere oggetto di un modello come in **figura 13C** dove è descritto come un'elica distesa

sulla superficie del tetramero ABCD. I linker che connettono il dominio catalitico e quello C-terminale delle subunità B e D si estendono dalla superficie concava, tanto da permettere ai domini C terminali di ancorare il DNA virale. Tale organizzazione va ben d'accordo con l'evidenza biochimica che il dominio C terminale lega una regione subterminale del DNA virale di circa cinque/dieci basi azotate dentro al sito di clivaggio⁶². Inoltre la disposizione della porzione legante il DNA virale lascia una piega basica tra i siti attivi che può accomodare il DNA target, come proposto per il Tn5. Comunque, dovrebbero essere necessari degli aggiustamenti alla struttura per legare sia il DNA virale che quello bersaglio; è anche richiesta la distorsione del DNA target in prossimità di ogni sito attivo per accomodarlo in questa piega⁶³. Parecchi modelli sono stati proposti per spiegare l'interazione tra il substrato ed un "multimero" di integrasi. Una caratteristica comune di questi modelli è che ogni terminale del DNA virale interagisce con il sito attivo, ed il dominio C terminale è dato dallo stesso dimero di integrasi. Contrariamente a quanto osservato nel modello mostrato in **figura 13C**, il sito attivo e il dominio C terminale sono dati da dimeri differenti.

Meccanismo d'azione dell'integrasi

Il processo di integrazione dell'HIV-1 comprende una serie di eventi molecolari che seguono il completamento della trascrizione inversa nel citoplasma della cellula infettata e che terminano con l'inizio della trascrizione del DNA provirale integrato. Il prodotto della trascrizione inversa è rappresentato da una doppia elica lineare di DNA con specifiche sequenze LTR (Long Terminal Repeat) alle estremità terminali. Queste ultime costituiscono il substrato per l'integrasi dell'HIV-1, l'enzima responsabile dell'inserimento del DNA virale nel cromosoma della cellula ospite ed essenziale per la replicazione del genoma virale. A questo scopo, l'integrasi catalizza una serie di reazioni consecutive riportate in **figura 14**.

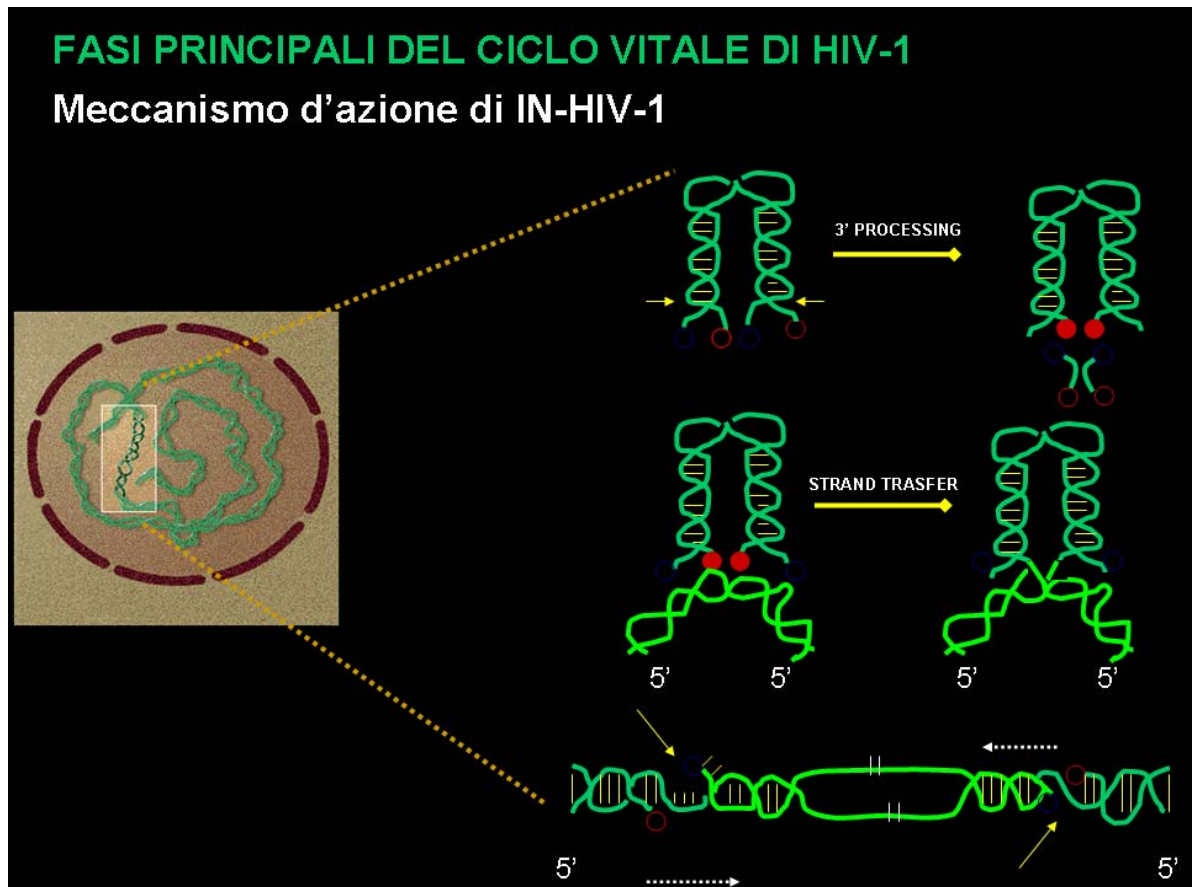


Figura 14. Meccanismo d'azione di IN

Nel primo step del processo di integrazione, chiamato “3’-processing”, l’enzima scinde due nucleotidi (GT) dall’estremità 3’-terminale del DNA virale (freccie gialle) all’interno della sequenza LTR, che contiene il motivo CA altamente conservato e produce degli idrossili 3’ reattivi ad entrambe le estremità del DNA virale (cerchi rossi pieni); le altre estremità 3’ idrossilate e 5’fosfato sono mostrate rispettivamente come punti contornati in rosso e blu rispettivamente). Questa reazione avviene nel citoplasma dove l’integrasi rimane

associata al DNA virale, formando il complesso di preintegrazione (PIC), una grande struttura molecolare contenente proteine strutturali, proteine virali accessorie e nucleocapside. Il complesso viene poi traslocato attraverso la membrana nucleare per accedere al DNA della cellula ospite mediante un meccanismo ATP dipendente, che consiste in un trasporto attivo attraverso i pori nucleari. Nel nucleo ha inizio quindi il secondo step, chiamato “strand transfer” in cui l’integrasi catalizza l’associazione tra DNA virale e DNA della cellula ospite. Dal punto di vista chimico avviene un attacco nucleofilo del gruppo 3’-OH terminale, formatosi nel 3’-processing, con concomitante transesterificazione del gruppo 5’-fosfato del DNA cellulare. Infine il processo di integrazione è completato dalla rimozione di un dinucleotide all’estremità 5’-terminale del DNA virale (freccie gialle in basso) e dal legame all’estremità 3’ del DNA cellulare: questo step è chiamato “gap-filling” o anche “5’-processing”. Sia il legame dell’integrasi al DNA cellulare che il 3’-processing e lo strand-transfer, possono essere valutati con saggi *in vitro*, usando integrasi ricombinante purificato (espresso in E. Coli) ed oligonucleotidi sintetici, in presenza di cationi bivalenti (Mg^{2+} e Mn^{2+}). Sequenze oligonucleotidiche specifiche rappresentano l’estremità U3 o U5 delle regioni LTR del DNA virale (substrato bersaglio). Questi studi dimostrano che le sequenze specifiche U3 e U5 del DNA virale sono indispensabili per la formazione di un complesso stabile enzimaticamente

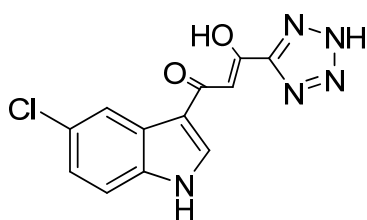
attivo; inoltre l'integrasi, per esplicare la propria attività catalitica, richiede la presenza dei metalli bivalenti Mn^{2+} e Mg^{2+} , il primo risulta più attivo *in vitro*.

In vitro, l'integrasi catalizza anche la reazione di disintegrazione: in questo processo, considerato come l'inverso dell'integrazione, l'integrasi scinde la porzione terminale del DNA virale e lega una porzione oligonucleotidica alle sequenze bersaglio vicine. Il significato biochimico di questa reazione durante il ciclo vitale del virus (*in vivo*) non è conosciuto, anche se è stato proposto un suo possibile ruolo nel 5'-processing.

Cenni di letteratura: progettazione del 5-CITEP e dei primi ligandi attivi contro l'integrasi di HIV-1

Il razionale terapeutico per lo sviluppo degli inibitori dell'integrasi è stato chiaro per molti anni, anche prima della recente convalida farmacologica dell'integrasi. Effettivamente, l'integrasi è essenziale per la replicazione virale e l'assenza di una controparte nella cellula ospite implica che gli inibitori dell'integrasi non interferiscano con i processi cellulari normali. Nuovi farmaci da studiare e nuovi target da esplorare inoltre sono necessari, perchè è noto che le terapie anti-HIV combinate sono molto più efficaci delle monoterapie: questo è il motivo per cui i regimi HAART si sono trasformati nello standard di cura per i pazienti di AIDS. La terapia di combinazione inoltre riduce la comparsa di virus farmaco-resistenti e di mutazioni multiple che sono necessarie al virus per instaurare la farmaco-resistenza. Probabilmente, gli inibitori dell'integrasi si trasformeranno in un potenziale aggiuntivo all'HAART o in una terapia di salvataggio per i pazienti resistenti ai farmaci anti-HIV attualmente disponibili. Negli ultimi anni sono stati condotti numerosi studi per determinare la struttura tridimensionale del sito catalitico dell'integrasi al fine di sviluppare un modello

per la progettazione di nuovi inibitori. In uno di questi studi, il dominio catalitico dell'integrasi è stato cocristallizzato con l'inibitore 5CITEP (1-(5-cloroindol-3-il)-3-idrossi-3-(2*H*-tetrazol-5-il)-propenone).



1

5-CITEP

L'inibitore si lega nella parte centrale del sito attivo dell'enzima, tra i tre residui acidi Asp-64, Asp-116 e Glu-152, fondamentali ai fini dell'attività catalitica. Inoltre la Gln-148 forma un legame idrogeno con l'azoto dell'anello indolico; il Glu-152 è alla distanza adatta per formare un legame idrogeno con l'ossidrilico enolico; tutti e quattro gli atomi d'azoto del tetrazolo si legano attraverso un legame idrogeno con Asn-155, Thr-66, Lys-159 e Lys-156. Non si osservano cambiamenti conformazionali nel complesso, eccetto che nella catena laterale del glutammato 152, molto flessibile. La presenza dell'inibitore, inoltre, non influenza il legame del magnesio ai due residui di aspartato. Studi di mutagenesi sito-specifica hanno permesso di identificare numerosi residui in

prossimità del sito attivo, importanti per il legame al DNA virale, come Lys-156, Lys-159, Gln-148 e Tyr-143. Alcuni di questi residui (Lys-156, Lys-159 e Gln-148) sono coinvolti anche nel legame con l'inibitore, come mostrato in **figura 15**. Si può quindi affermare che il legame dell'inibitore mima effettivamente l'interazione integrasi-DNA virale. Questa osservazione è confermata dal fatto che la distanza tra l'indolo e il tetrazolo nell'inibitore può essere approssimata a quella esistente tra due nucleotidi.

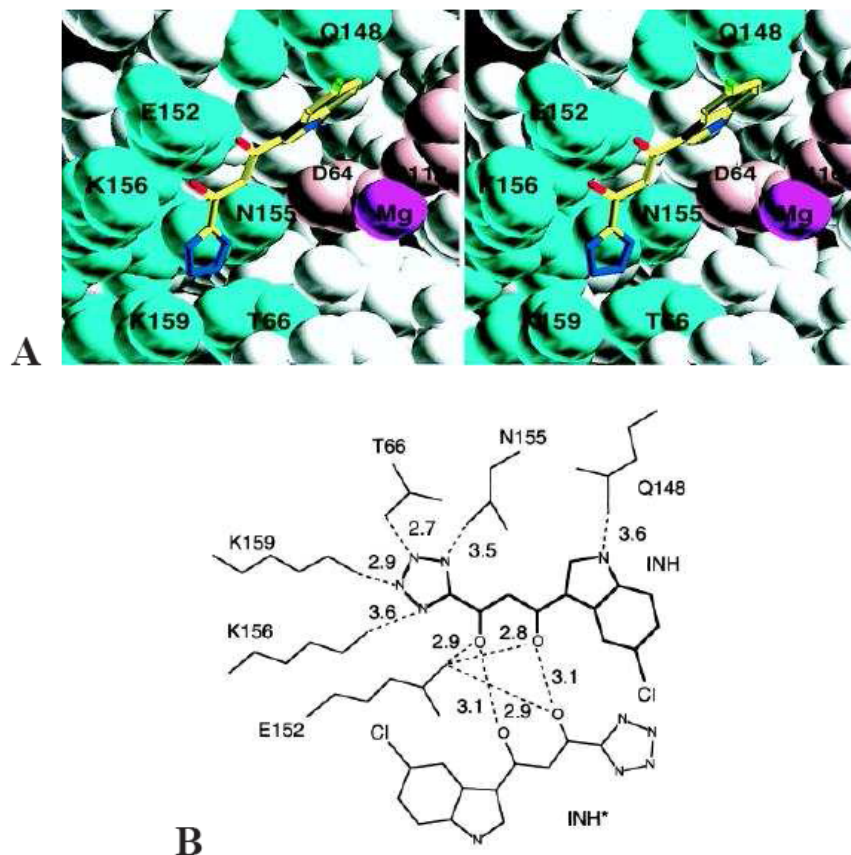


Figura 15. Immagine delle interazioni tra l'inibitore e l'enzima. Le catene laterali dell'enzima che interagiscono sono mostrate in turchese, il magnesio in viola ed i residui Asp-64 ed Asp-116 in rosa (A). Rappresentazione schematica. Distanze corrispondenti a potenziali legami idrogeno (B).

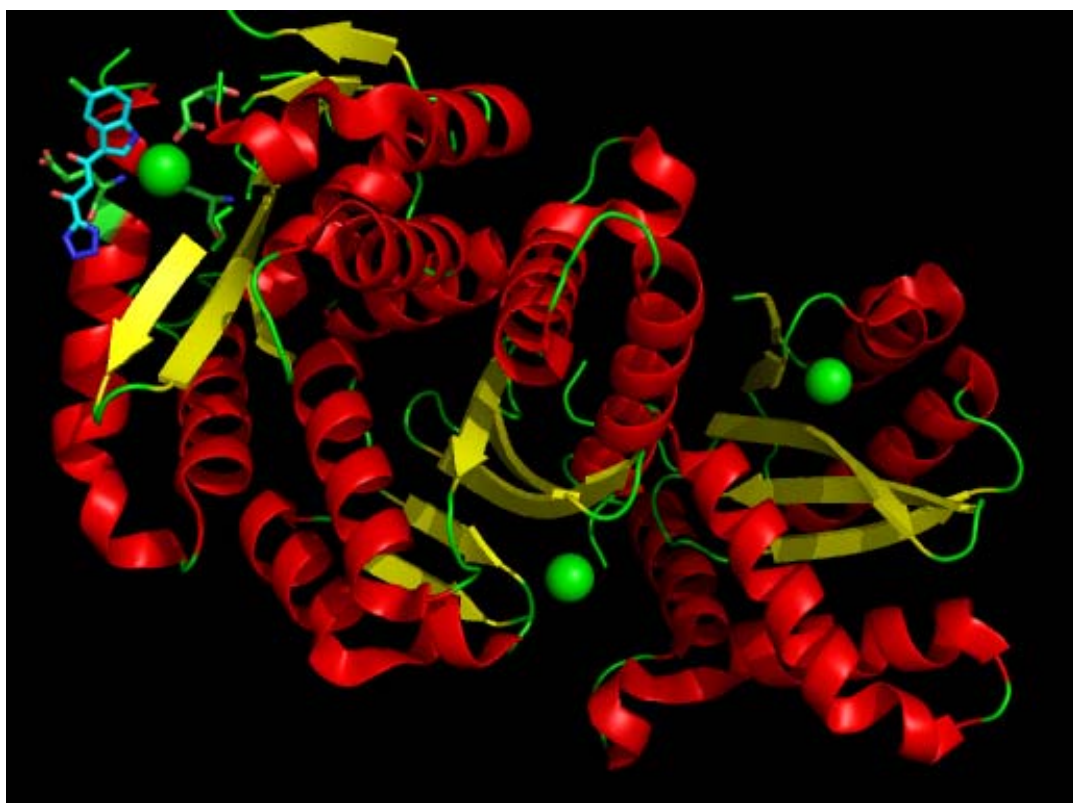
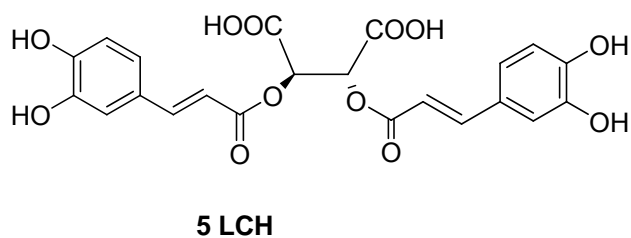
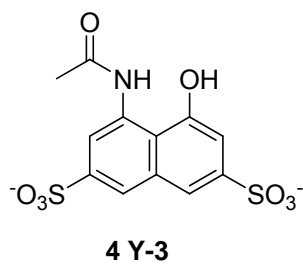
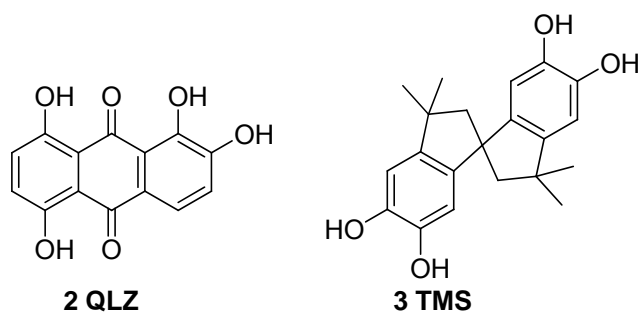


Figura 16. Rappresentazione del dominio catalitico cocristallizzato con l'inibitore 5CITEP (codice PDB: 1QS4).

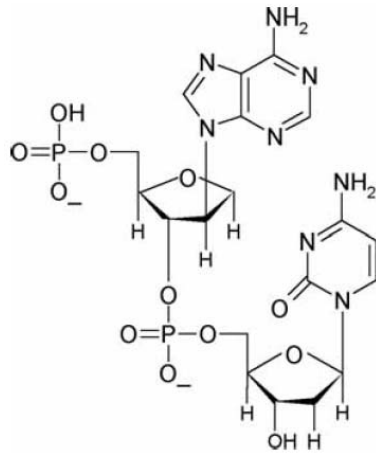
Il 5CITEP può essere in effetti considerato un nucleotide, in cui il tetrazolo rimpiazza il fosfato, la porzione chetoenolica mima lo zucchero, ed il cloroindolo sostituisce la base purinica⁶¹.

Recentemente sono stati condotti degli studi di dinamica molecolare sul complesso 5CITEP-integrasi che hanno permesso di risalire alle esatte interazioni tra l'integrasi e il ligando. Durante la simulazione di dinamica, il 5CITEP ha mostrato una considerevole mobilità all'interno del sito attivo, spostandosi dalla

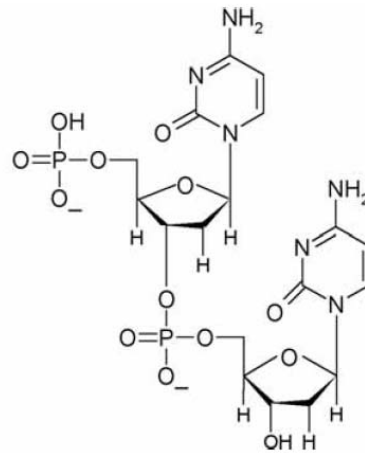
posizione iniziale assunta nella struttura cristallizzata. Inoltre, è stata rilevata la presenza di una molecola d'acqua nel sito catalitico che sembra favorire il legame tra inibitore e integrasi: essa forma, infatti, un legame idrogeno tra il 5CITEP e l'integrasi, agendo da ponte di connessione⁶². Sulla base della struttura cristallina del complesso tra il 5CITEP ed il sito attivo dell'integrasi, sono stati realizzati modelli strutturali per l'interazione di vari inibitori attraverso il "computational docking". Gli inibitori analizzati sono: 1,2,5,8-tetraidrossiantrachinone (QLZ), 3,3,3',3'-tetrametil-1,1'-spirobis(indan)-5,5',6,6'-tetrololo (TMS), acido L-cicorico (LCH), acido 4-acetilamino-5-idrossinaftalen-2,7-disolfonico (Y-3), composti 2-5.



Pertanto anche mononucleotidi e dinucleotidi come sistemi modello per l'analisi delle interazioni DNA-proteine. In uno di questi studi, due dinucleotidi modificati, composti **6-7**, sono stati inseriti nel sito attivo dell'integrasi, utilizzando il programma "Autodock".



6 pispApdC



7 pdCpisodU

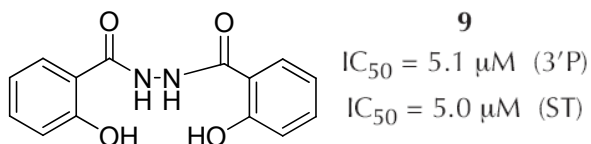
Questi dinucleotidi, in contrasto con la loro conformazione ammassata, assumono invece una conformazione distesa all'interno del sito catalitico. Ciò è dovuto alle estese interazioni favorevoli che si instaurano con i residui amminoacidici del sito attivo, che facilitano la formazione di un legame stabile tra i dinucleotidi e l'integrasi. Sono stati, inoltre, identificati i siti di interazione con il DNA virale e il DNA cellulare: alcuni di questi siti sovrapponibili con il dominio catalitico dell'integrasi interagiscono con il DNA virale, mentre un altro gruppo di siti, lontani dai primi, vanno ad interagire con il DNA cellulare. L'interazione con il DNA virale risulta quindi sequenza-dipendente, mentre non lo è quella con il DNA cellulare. Naturalmente è necessario lavorare ancora molto per migliorare le conoscenze sulle interazioni DNA-integrasi, indispensabili per lo sviluppo di nuovi inibitori selettivi⁶⁰.

Questi dinucleotidi, in contrasto con la loro conformazione ammassata, assumono invece una conformazione distesa all'interno del sito catalitico. Ciò è dovuto alle estese interazioni favorevoli che si instaurano con i residui amminoacidici del sito attivo, che facilitano la formazione di un legame stabile tra i dinucleotidi e l'integrasi. Sono stati, inoltre, identificati i siti di interazione con il DNA virale e il DNA cellulare: alcuni di questi siti sovrapponibili con il dominio catalitico dell'integrasi, interagiscono con il DNA virale, mentre un altro gruppo di siti, lontani dai primi, vanno ad interagire con il DNA cellulare. L'interazione con il DNA virale risulta quindi sequenza-dipendente, mentre non lo è quella con il DNA cellulare. Naturalmente è necessario lavorare ancora molto per migliorare le conoscenze sulle interazioni DNA-integrasi, indispensabili per lo sviluppo di nuovi inibitori selettivi.

Mercaptosalicilidrazidi

Le mercaptosalicilidrazidi sono composti che hanno un'attività inibitoria legata alla loro capacità di legare i metalli bivalenti. Derivano dalla salicilidrazide, composto **9** e composto **10**, rispetto alla quale risulta meno tossico con un'attività inibitoria selettiva nei confronti dell'integrasi.

Salicilidrazide

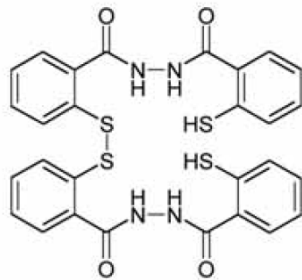


La salicilidrazide inibisce sia il 3'-processing che lo strand transfer, con una IC_{50} pari a 2.0 μM e 0.7 μM rispettivamente, in presenza del Mn^{2+} , mentre risulta totalmente inattiva in presenza del Mg^{2+} .

Le mercaptosalicilidrazidi derivate come nel caso del composto **10**, invece, inibiscono l'integrasi in presenza sia dell'uno che dell'altro metallo. La loro

attività inibitoria consiste nel bloccare il legame dell'integrasi al DNA substrato, inibendo quindi la formazione del complesso INDNA.

Le mercaptosalicilidrazidi possiedono infatti uno o due mercaptogruppi liberi che legano la Cys-65 nel sito attivo dell'integrasi e chelano il metallo bivalente con la funzione dichetonica.⁶⁴



10

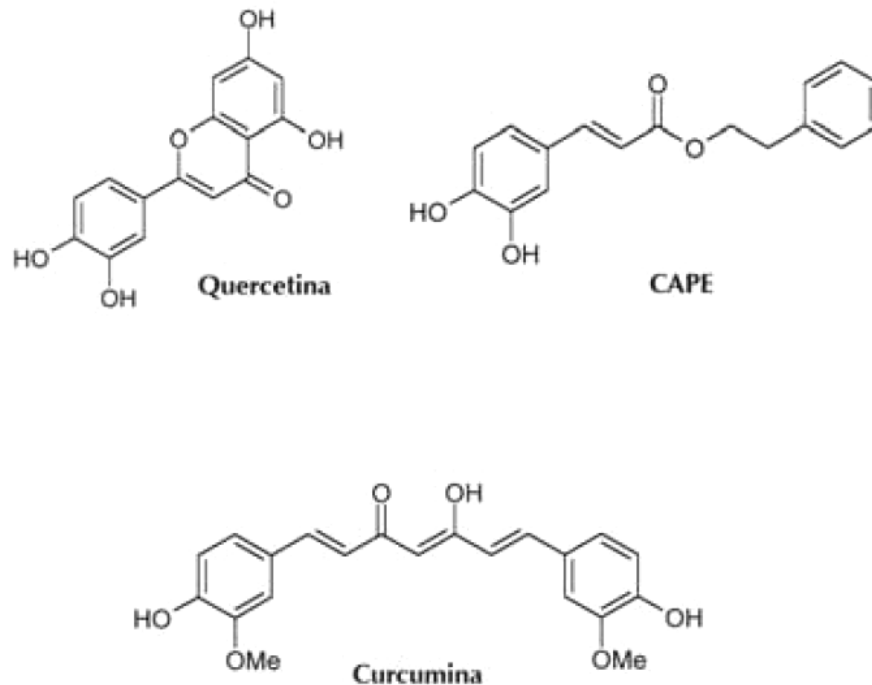
$IC_{50} = 3.2 \mu M$ (3'P)

$IC_{50} = 3.7 \mu M$ (ST)

Composti ossidrilati

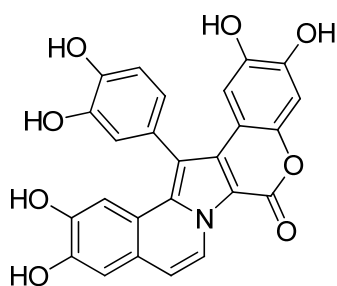
Tra i composti ossidrilati, uno dei primi a mostrare attività antiintegrasi è stato l'acido aurintricarbossilico. Dopo di questo, altre molecole polioidrossilate, sia naturali che sintetiche, si sono rivelate inibitori dell'integrasi: la quercetina, l'estere fenilico dell'acido caffeico (CAPE) e le ammidi ad esso correlate, bis-

catecoli come la curcumina, la tirfostina, l'acido L-cicorico (L-CCA) e l'acido 3,5-dicaffeoilchinico (3,5-DCQA).



Sfortunatamente, solo pochi di essi hanno mostrato attività nei saggi cellulari. Le uniche eccezioni sono l'acido L-cicorico ed alcuni acidi dicaffeoilchinici che, oltre ad essere inibitori potenti e specifici dell'integrasi, inibiscono la replicazione del virus HIV- 1 nelle cellule infettate a concentrazioni non citotossiche. Moderatamente attivo nei saggi cellulari si è mostrato anche un altro inibitore poliossidrilato, la curcumina, estratta dai rizomi della *Curcuma longa*, che presenta anche attività antinfiammatoria e antitumorale. La modesta attività presentata dalla curcumina è probabilmente dovuta alla sostituzione del

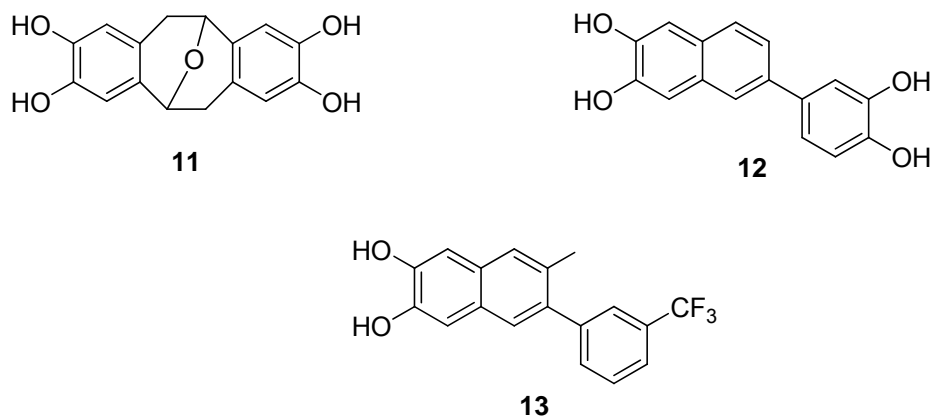
gruppo catecolico con il sistema *o*-metossiidrossi: sembra, infatti, che il catecolo sia indispensabile per l'inibizione dell'attività enzimatica⁶⁵⁻⁶⁶. Tra i composti ossidrilati di origine naturale che contengono il sistema catecolico, vanno ricordati gli analoghi delle lamellarine, un gruppo di alcaloidi esaciclici, isolate da molluschi appartenenti al genere *Lamellaria*. Il composto più attivo di questa serie di derivati è la lamellarina H, che presenta una IC₅₀ per lo strand transfer pari a 1.3 μM.



Lamellarina H

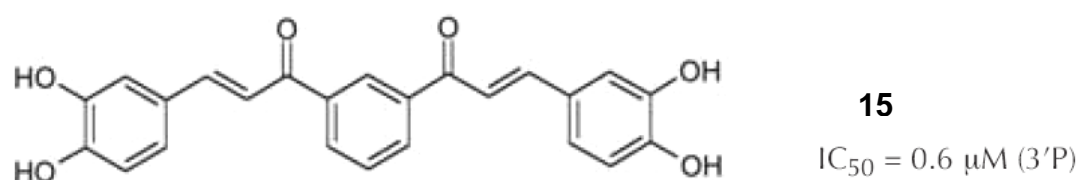
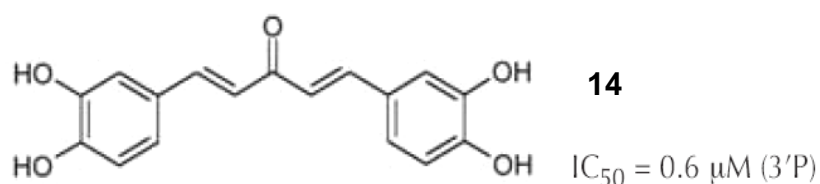
Sebbene questo composto inibisca l'integrasi a basse concentrazioni micromolari, è risultato un inibitore non-selettivo e mostra una elevata citotossicità (CC₅₀ = 5.7 μM). Sulla base di tali risultati sono stati quindi sintetizzati diversi derivati di origine catecolica ed è stata valutata la loro attività

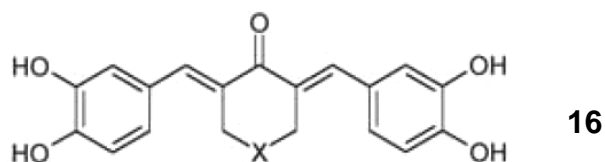
inibitoria. Quelli che hanno mostrato una maggiore attività anti-integrasica sono i seguenti composti 11-13:



Essi inibiscono il 3'-processing con una IC_{50} pari a 1.3 μM (11), 2.1 μM (12), 13 μM (13) e mostrano una certa attività antivirale con una EC_{50} pari a 98 μM (11), 9.9 μM (12) e 5.7 μM (13). Non è stato ancora chiarito completamente il meccanismo molecolare di interazione tra l'enzima e questi composti, che potrebbe consistere nella coordinazione del gruppo catecolico con i cationi bivalenti (Mn^{2+} e Mg^{2+}), necessari per la catalisi enzimatica, oppure nella formazione di legami idrogeno, in cui gli ossidrili si comportano da donatori, con siti specifici del dominio catalitico dell'enzima. Bisogna comunque considerare che il gruppo catecolico può subire, *in vivo*, un processo di ossidazione che porta alla formazione di chinoni, specie molto reattive e

facilmente attaccate da nucleofili endogeni. Queste reazioni sono alla base della tossicità dei derivati catecolici, giustificando la loro scarsa attività nei saggi cellulari a concentrazioni sub citotossiche. Nel tentativo di migliorare l'attività dei derivati catecolici, sono stati sintetizzati diversi composti **14-16**, caratterizzati da una struttura cinnamoilica e da una disposizione *syn* dei doppi legami C=C e C=O. Questa costrizione geometrica o conformazionale, osservata nella curcumina (composto poliossidrilato naturale attivo come inibitore dell'integrasi), causa infatti un aumento di attività sulle cellule infettate. Sebbene questi composti mostrino la caratteristica citotossicità del sistema catecolico, alcuni derivati sono risultati potenti inibitori dell'integrasi nei saggi enzimatici

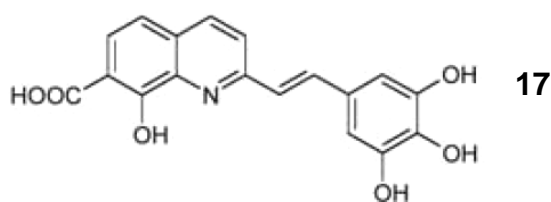




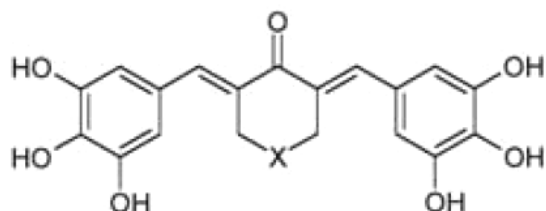
X = O, S, NH, NR, CH₂

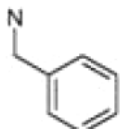
I più attivi di questo gruppo di inibitori sono risultati i derivati del cicloesanone (**16**), con valori di IC₅₀ tra 0.2 e 0.9 μM. Inoltre, analisi 3D QSAR CoMFA (Comparative Molecular Field Analysis) e CoMSIA (Comparative Molecular Similarity Indices Analysis) e simulazioni di docking, condotti su questi composti, hanno documentato l'importanza del legame idrogeno di questi inibitori nell'interazione con il sito attivo dell'enzima⁶⁷.

L'attività mostrata in particolare da alcune stililchinoline sostituite, tra cui il composto **17**, capaci di inibire l'integrasi e bloccare la replicazione dell'HIV-1 in colture cellulari, ha evidenziato la possibilità di sostituire il gruppo 3,4-diidrossicinnamoile con il gruppo 3,4,5-triidrossicinnamoile.



Questi composti conservano infatti la capacità di chelare i cationi bivalenti e formare legami idrogeno e, allo stesso tempo, mostrano una minore citotossicità, peculiare del sistema catecolico. Sono stati perciò progettati e sintetizzati nuovi inibitori a struttura triidrossicinnamoilica, analoghi dei derivati catecolici precedentemente descritti

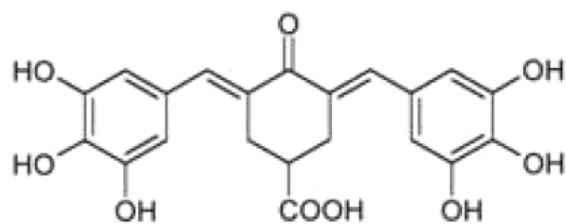


X	IC ₅₀ (3P)	IC ₅₀ (ST)	EC ₅₀	CC ₅₀
S	0.7 μM	0.5 μM	20 μM	100 μM
O	6.0 μM	9.0 μM	34 μM	115 μM
	0.7 μM	1.1 μM	30 μM	117 μM

Rispetto ai diidrossiderivati, presentano una più potente attività inibitoria nei confronti dell'integrasi, sono meno tossici ed, inoltre, sono attivi anche nei saggi cellulari, rivelandosi in grado di inibire la moltiplicazione virale a concentrazioni inferiori a quelle citotossiche. Tutto ciò sembra dovuto

all'introduzione del terzo ossidrilico, che riduce il processo di alchilazione delle proteine cellulari ed è posizionato in una porzione della molecola pienamente coinvolta nella formazione del legame idrogeno con il bersaglio biologico⁶⁷.

Inoltre, la correlazione tra la presenza di gruppi carbossilici e l'attività nei confronti dell'integrasi e sulla moltiplicazione dell'HIV-1 nelle cellule infettate, osservata nell'acido L-cicorico e nelle stilichinoline, ha giustificato l'introduzione della funzione carbossilica nei triidrossicinnamoilderivati composti 17-19. Grazie a questo, è stata aumentata di molto la potenza degli inibitori, probabilmente per le interazioni addizionali del gruppo carbossilico con la Lys-156 e la Lys-159, presenti nel sito catalitico dell'HIV-1 integrasi⁶⁶⁻⁶⁷.

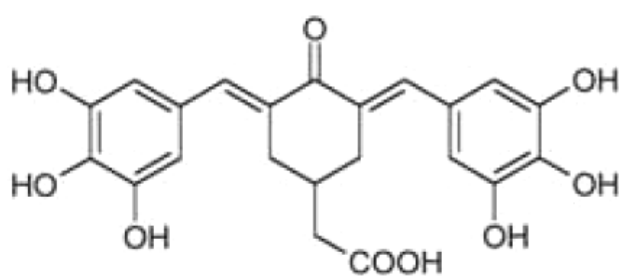


17

$IC_{50} = 0.2 \mu M$ (3'P)

$IC_{50} = 0.3 \mu M$ (ST)

$EC_{50} = 4 \mu M$

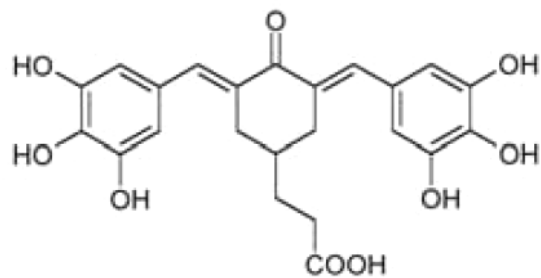


18

$IC_{50} = 0.2 \mu M$ (3'P)

$IC_{50} = 0.2 \mu M$ (ST)

$EC_{50} = 2 \mu M$



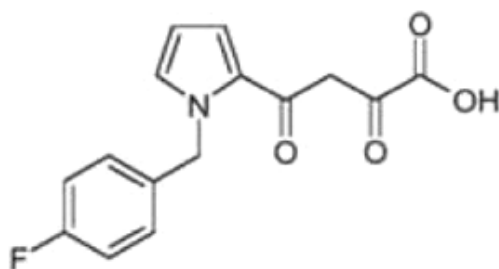
19

La ricerca di nuovi composti poliossidrilati come inibitori dell'integrasi è stata abbandonata in seguito alla scoperta che l'attività dell'acido L-cicorico e dei suoi analoghi nei confronti dell'HIV *in vivo* non è legata all'inibizione dell'integrasi, ma al blocco della "viral entry". Tali derivati agiscono quindi principalmente sul legame della cellula alla proteina virale *gp120*.

Arildichetoacidi

Il 5CITEP composto 1 è un analogo strutturale dei dichetoacidi, in esso, infatti, la porzione dichetonica connette un tetrazolo (isostero del residuo acido) ad un anello idrofobico. Il 5CITEP lega l'integrasi anche in assenza del DNA substrato, senza provocare sostanziali modificazioni conformazionali nella proteina⁶⁸.

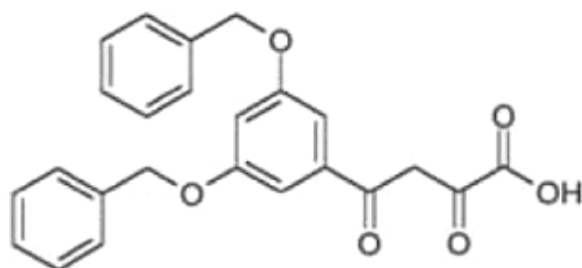
Gli arildichetoacidi costituiscono una classe di composti in grado di inibire in modo specifico l'integrasi del virus HIV-1. I lead-compounds, per questa classe di inibitori, sono i dichetoacidi L-731,988 (20) e L-708,906 (21).



20 L-731,988

$IC_{50} = 0.17 \mu\text{M}$ (3'P)

$IC_{50} = 6 \mu\text{M}$ (ST)



21 L-708,906

$IC_{50} = > 1000 \mu\text{M}$ (3'P)

$IC_{50} = 0.48 \mu\text{M}$ (ST)

I dichetoacidi possono essere considerati degli inibitori selettivi dello strand transfer in quanto impediscono la replicazione dell'HIV-1, bloccando tale step dell'integrazione, a concentrazioni significativamente più basse di quelle richieste per inibire il 3'-processing. Sebbene siano richiesti gli stessi residui all'interno del sito attivo sia per il 3'-processing che per lo strand transfer, i di chetoacidi sembrano in grado di discriminare tra le due funzioni catalitiche. L'associazione dell'integrasi con le sequenze LTR sembra portare alla formazione di un complesso attivo e stabile. È stato dimostrato che il composto **L-1,988** riconosce l'integrasi esclusivamente nel contesto di una proteina cataliticamente attiva,

associata all'estremità del DNA virale. Esso si lega al sito attivo dell'integrasi ed inibisce lo strand transfer competendo con il DNA substrato⁶⁹. Gli studi condotti sui composti **20-21** hanno dimostrato che essi, oltre ad essere molto attivi nei saggi enzimatici, sono anche dei potenti inibitori del complesso di preintegrazione. Nei saggi cellulari, inibiscono la replicazione dell'HIV-1 con una EC₅₀ rispettivamente di 1 e 2 μM. Sono inoltre attivi contro ceppi mutanti di HIV-1, resistenti agli inibitori della proteasi e della trascrittasi inversa⁶⁹.

Recentemente sono stati sintetizzati una serie di derivati analoghi a quelli descritti in precedenza, basandosi sulla struttura del 5CITEP e dell'L-708,906, e ne è stata valutata la loro attività antivirale e anti-integrasi. Queste molecole presentano come motivo comune un linker β-dichetopropanico, ma differiscono sostanzialmente nella porzione aromatica e nel gruppo acido. (figura 17).

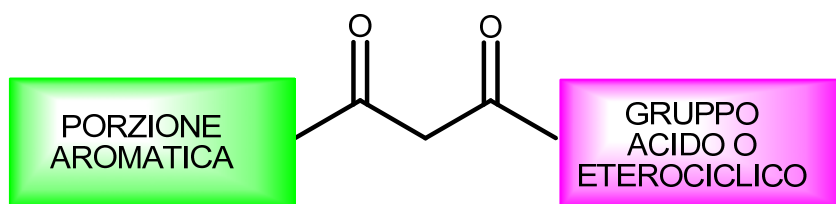
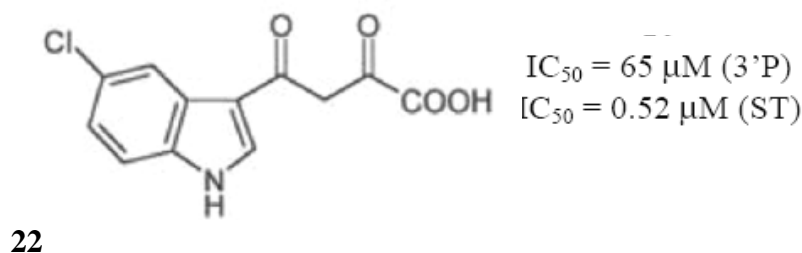


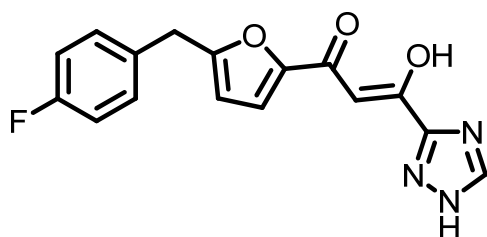
Figura 17. Linker β-dichetopropanico

Per osservare l'influenza della funzione acida sull'inibizione dell'HIV-1, è stato testato un analogo del 5CITEP, nel quale il tetrazolo è stato sostituito con un gruppo carbossilico composto **22**. Questo composto mostra un'attività inibitoria dell'integrasi paragonabile a quella del 5CITEP, confermando che il tetrazolo ed il gruppo carbossilico sono bioisosteri.



Dal momento che il composto **22** e L-708,906 (**21**) differiscono unicamente nella porzione aromatica, questa parte della molecola è ritenuta responsabile della differenza osservata nell'inibizione del 3'-processing e della selettività dei dichetoacidi per lo strand transfer^{70,71}. Inoltre è stato osservato che per il legame del dichetoacido con 'integrasi sono necessari gli ioni bivalenti e che i derivati isosterici dei dichetoacidi possono esibire metallo-specificità. Infatti i composti che presentano la funzione carbossilica legano sia il Mn^{2+} che il Mg^{2+} , mentre quelli che presentano anelli eterociclici contenenti atomi di azoto hanno

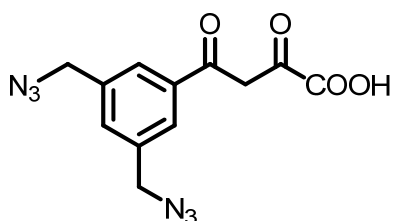
alta affinità per il Mn^{2+} . A questa classe di inibitori appartiene anche il composto S-1360 composto **23**, sintetizzato dalla Shionogi-GlaxoSmithKline Pharmaceuticals.



23 S-1360

L'S-1360 è un potente inibitore dell'integrasi di HIV-1 *in vitro*, con valori di IC_{50} pari a 20 nM, e mostra una notevole attività antivirale (EC_{50} pari a 140 nM), con effetto citotossico solo ad elevata concentrazione ($CC_{50} = 110 \mu M$). Il composto è attivo contro ceppi di HIV-1 con tropismo verso cellule T e macrofagi ed inibisce la replicazione di virus HIV-1 resistenti agli inibitori della trascrittasi inversa e della proteasi. L'S-1360 ha dimostrato un'attività sinergica *in vitro* se usato in combinazione con vari inibitori nucleosidici e non-nucleosidici della trascrittasi inversa e in combinazione con inibitori della proteasi, come azidotimidina (AZT), lamivudina, nevirapina e nelfinavir⁷². Nello studio di nuovi inibitori dell'integrasi, sono stati recentemente introdotti gli

azido-aril- β -dichetoacidi, in cui il gruppo N_3 rimpiazza un arile degli inibitori tradizionali. Il composto **24**, mostra un'attività inibitoria nei confronti dell'integrasi paragonabile a quella dell'**L-708,906** di cui è strutturalmente analogo, ed un'attività antivirale in colture cellulari, ma soprattutto mostra una bassa citotossicità ($CC_{50} > 50 \mu M$).



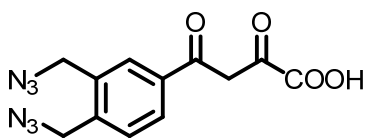
24

L'inaspettata capacità del gruppo azido di sostituire la funzione arilica, ha indotto ad esaminare questa classe di composti ed alla sintesi di numerosi derivati, tra cui il più attivo è il composto **25**.

$$IC_{50} = >100 \mu M (3'P)$$

$$IC_{50} = 0.26 \mu M (ST)$$

$$EC_{50} = 14.3 \mu M$$



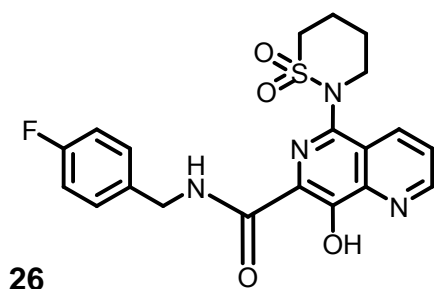
25

L'analogo 3,4-disostituito (25) risulta approssimativamente dieci volte più attivo del composto 24, 3,5-disostituito. Questo aumento di attività può essere potenzialmente attribuito all'orientamento preferenziale del gruppo azido, indotta dalla *orto*-disostituzione. Inoltre dall'analisi di mono-azido derivati, risulta che entrambi i gruppi sono importanti ai fini dell'attività inibitoria. Di notevole importanza sono anche le proprietà steriche ed elettroniche: infatti le azidi sono stericamente caratterizzate da una geometria lineare e presentano una regione terminale con un potenziale elettrostatico negativo che interagisce con gli ioni metallici bivalenti all'interno del sito attivo dell'enzima⁷³. La Merck Research Laboratories ha proposto negli ultimi anni una nuova classe di inibitori dell'integrasi: le naftiridin-carbossamidi, derivati strutturalmente dai dichetoacidi. Il composto L-870,810 (26) è attualmente sottoposto alle ultime fasi della sperimentazione clinica.

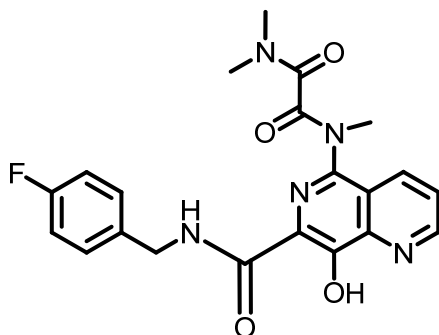
IC₅₀ = >0.01 μM

EC₅₀ = 0.004 μM

CC₅₀ = > 5 μM



Il composto **L-870,810** è un inibitore potente e selettivo dello strand transfer, e mostra un buon profilo farmacocinetico ed una maggiore biodisponibilità nei confronti dei dichetoacidi precedentemente riportati. **L-870,810** inibisce l'attività integrasica ST *in vitro* con valori di IC₅₀ pari a 0.01 μM e presenta attività antivirale con valori di EC₉₅ pari a 0.019 μM nei saggi cellulari. Ulteriori sforzi di sintesi e di ottimizzazione condotti dal gruppo Merck hanno recentemente condotto al derivato **L-870,812**, composto **27**.



27 L-870,812

Esso è un promettente lead per lo sviluppo terapeutico grazie alla sua selettività anti-integrasi, alla sua potenza anti-virale e al suo profilo farmacocinetico favorevole (biodisponibilità orale > 60% ed emivita di circa 5 ore nei macachi rhesus). L'attività antivirale di **L-870,812** è stata recentemente dimostrata, il composto è infatti in grado di abbattere la carica virale così come le infezioni croniche nei macachi rhesus infettati con il Simian Immunodeficiency Virus. Il successo clinico dei composti S-1360 e L-870,810 potrebbe rappresentare una scoperta importante nello sviluppo dei dichetoacidi e dei loro analoghi come inibitori dell'integrasi. L'inibizione selettiva dello strand transfer a concentrazioni nanomolari è una caratteristica degli arildichetoacidi (DKAs), anche se il CAPE è stato il primo inibitore dello strand transfer ad essere riportato in letteratura. Tuttavia, i benefici di tale inibizione selettiva nel trattamento dell'infezione da HIV-1 sono emersi soltanto con la descrizione dei

DKAs più potenti. Di conseguenza, gli arildichetoacidi sono i primi inibitori dell'integrasi a fornire la prova del concetto che l'inibizione selettiva dello strand transfer può far effettivamente diminuire l'infettività da HIV. Un metallo bivalente è richiesto come cofattore per il legame del DKA all'integrasi e l'inibizione dell'integrasi può essere dipendente da un catione divalente. Per esempio, la IC_{50} di **L-708,906** per il 3'-processing è 2.5 μM in Mg^{2+} e 22 μM in Mn^{2+} , mentre la relativa IC_{50} per lo strand transfer è paragonabile con entrambi i metalli (selettività = 0.06). La selettività per lo strand transfer può anche essere metallo-bivalente dipendente. Sia l'**L-708,906** che il 5CITEP hanno inibizione approssimativamente dieci volte più alta in presenza di Mn^{2+} rispetto al Mg^{2+} . Inoltre, il 5CITEP è un inibitore molto più efficace in Mn^{2+} che in Mg^{2+} (uno-due ordini di grandezza più potente). Nonostante la loro selettività per lo strand transfer, i DKAs possono anche inibire il 3'-processing, anche se a volte sono necessarie concentrazioni 30-70 volte maggiori. Gli studi SAR hanno rivelato il contributo dato all'inibizione dell'integrasi dalle porzioni acide ed aromatiche dei dichetoacidi. L'attività relativamente scarsa del 5CITEP in Mg^{2+} è dovuta alla funzione tetrazolica; ciò è dimostrato dal fatto che, rimpiazzando il tetrazolo con un carbossilato marcato, aumenta l'attività inibitoria del composto ibrido in Mg^{2+} . Il carbossilato, quindi, potrebbe essere importante per la chelazione del metallo. La porzione carbossilica, tuttavia, non è indispensabile per il legame al

complesso dell'integrasi. Gli studi SAR rivelano che la parte aromatica è cruciale per la potenza e per la selettività nei confronti dello strand transfer. La parte aromatica può ospitare una vasta gamma di sostituenti, compreso l'azido-gruppo. La porzione azidica dei di chetoacidi contribuisce a ridurre la citotossicità e potrebbe avere un ruolo diretto nella chelazione del metallo. Il dichetogruppo o il gruppo funzionale β -idrossichetonico sono noti per la loro capacità di chelare i metalli e l'inibizione metallo-dipendente dei DKAs è stata interpretata come l'indicazione di un'interazione diretta di questi farmaci con il metallo bivalente nel sito catalitico dell'enzima. La coordinazione del metallo è stata inoltre importante per predire un modello della tasca catalitica dell'integrasi e del sito legante i DKAs. Il legame molecolare dei DKAs ai complessi di integrazione è stato al centro di un'accesa ricerca a causa dell'importanza dei DKAs e dei farmaci DKA-simili come lead compound antivirali e per il loro particolare meccanismo d'azione. L'alta selettività dei DKAs per lo strand transfer porta ad un modello in cui i due siti catalitici sono organizzati intorno ai tre residui del motivo DDE ed al metallo bivalente all'interno dei complessi integrasi-DNA (figura 18).

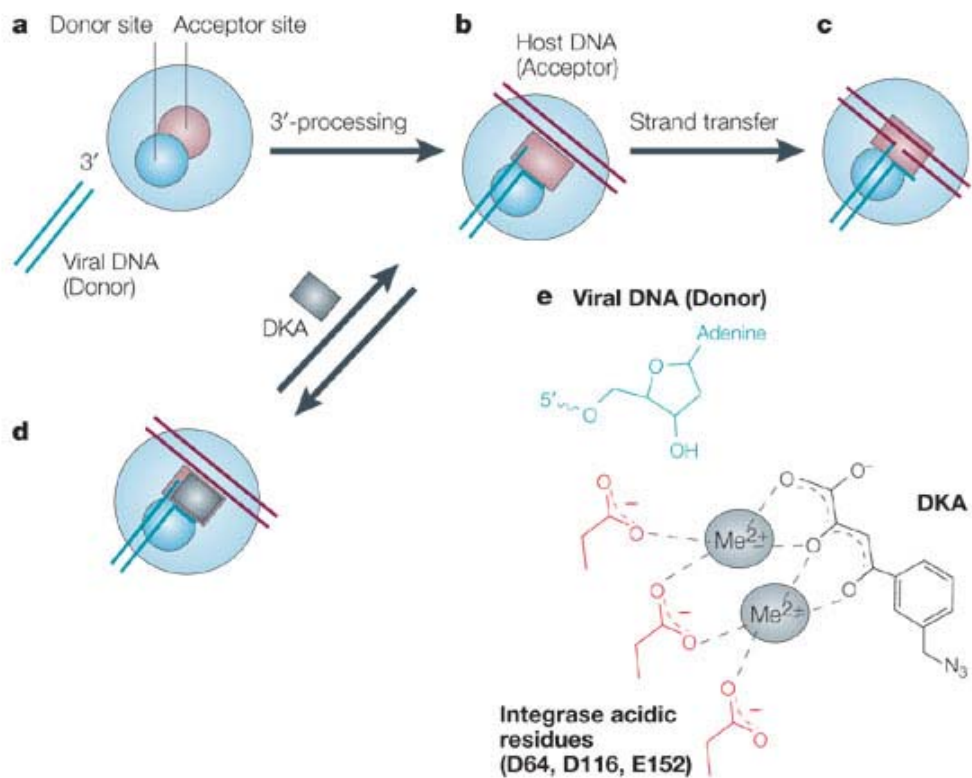


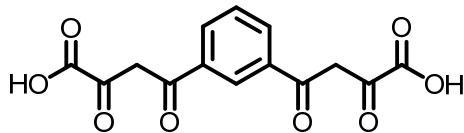
Figura 18 Modello di legame dei DKA nel sito attivo dell'integrasi

In questo modello, il sito del DNA donatore (donor site, **figura 18a**) lega l'estremità del DNA donatore (virale) e catalizza il 3'-processing. Di conseguenza, l'integrasi subisce un mutamento strutturale che permette il legame del DNA (cromosomico) dell'acceptor nel sito dell'acceptor (acceptor site, **figura 18b**) per lo strand transfer. I DKAs si legherebbero selettivamente ad una unica conformazione del sito acceptor dopo il legame del DNA virale ed il 3'-

processing. Solo allora si avrebbe un cambiamento conformazionale del sito accettore per i DKA-ligandi. Secondo questo schema, la coordinazione bivalente del metallo sarebbe cruciale per il legame del DKA al sito accettore. Questo modello è sostenuto dalle analisi con composti marcati, che ha dimostrato il legame dei DKAs ad un intermedio del PIC in presenza di Mg^{2+} o di Mn^{2+} . Il legame dei DKAs richiede un'integrasi funzionale, poichè i mutanti e gli enzimi cataliticamente inattivi non sono riusciti a legare i DKAs. I farmaci legati richiedono invece le estremità virali del DNA, poiché un DNA non specifico non riesce ad essere legato. Per concludere, i DKAs non riescono a legarsi efficacemente al DNA-target nella reazione di strand transfer. Le caratteristiche strutturali del sito legante i DKAs possono anche essere dedotte dalle mappe dei residui amminoacidici connessi con la DKA-resistenza nella struttura cristallina dell'integrasi. I mutanti dell'integrasi sono stati selezionati da DKA con mutazioni nel CCD (ed una nel CTD). I residui amminoacidici all'interno del CCD che conferiscono resistenza sono raggruppati intorno al sito catalitico definito dai residui del DDE. Questo gruppo coincide con il gruppo dei residui interessati al legame del DNA. Un residuo (Q148) è comune ad entrambi i gruppi e sembra partecipare al corretto posizionamento del DNA virale sul DNA accettore prima dello strand transfer. Il gruppo di residui coinvolti sia nella resistenza del farmaco (e quindi probabilmente nel legame dei DKAs) che nel

contatto del DNA con la triade catalitica che coordina il metallo, aumenta la probabilità che i DKAs si leghino all'interfaccia del complesso ternario DNA virale (donatore)-integrasi-metallo. Il legame dei DKAs avverrebbe specificamente nello step del ciclo catalitico che precede lo switching del complesso dal 3'-processing allo strand transfer (cioè quando è assemblato il PIC). Tale legame stabilizzerebbe l'intermedio derivante dal 3'-processing ed impedirebbe lo strand transfer. Recentemente è stato proposto che questo meccanismo di inibizione fosse un meccanismo comune anche per i composti naturali che interferiscono con i complessi macromolecolari proteina-DNA (quali il complesso camptotecinatopoisomerasi I, il dexrazoxano e le antracicline per il complesso topoisomerasi II DNA o i complessi antibiotico-ribosoma) o con i complessi proteina-proteina (come il complesso tra la brefeldina A e l'Arf-GEP o il complesso colchicina- α,β -tubulina). Si può quindi proporre che DKAs vengano classificati come "inibitori interfacciali" dei complessi macromolecolari. Un ulteriore approfondimento riguardo il meccanismo d'azione dei DKAs deriva dall'osservazione che si ha un marcato aumento di attività inibitoria del 3'-processing ed una riduzione di selettività 3'P/ST nel caso dei di chetoacidi bifunzionali (BDKAs). Il composto **28**, analogo dell' **L-708,906**, in cui la porzione aromatica è stata modificata attraverso l'introduzione di una seconda catena dichetoacida, conserva l'attività sullo strand transfer, ma inibisce il 3'-

processing con valori di IC_{50} inferiori a 10 μM , perdendo la selettività nei confronti dello strand transfer.



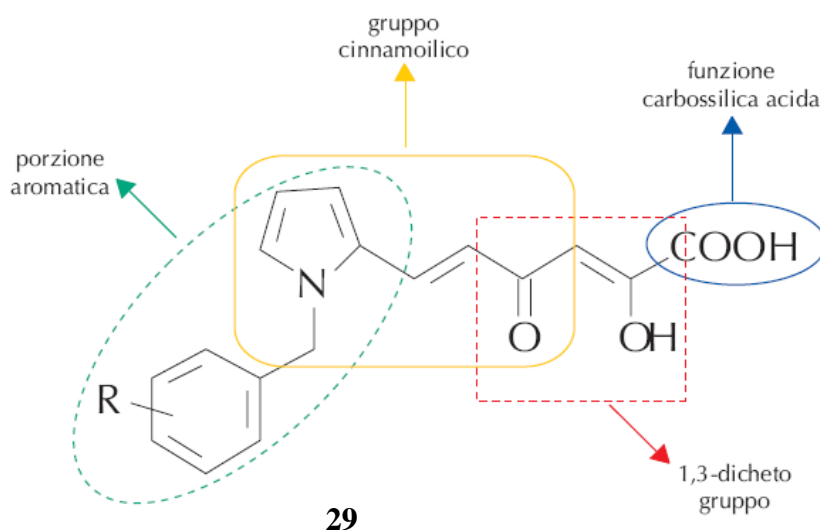
28

$IC_{50} = 7.8 \mu M$ (3'P)

$IC_{50} = 1.83 \mu M$ (ST)

La differente selettività di questi composti indica probabilmente un distinto meccanismo d'azione e una diversa interazione con l'enzima⁷⁶⁻⁷⁷. È stato infatti ipotizzato che i dichetoacidi monofunzionali, come l'**L-708,906**, si legano al sito, detto accettore, dove l'enzima catalizza l'attacco nucleofilo dell'estremità 3'-OH terminale del DNA virale (donatore), formando un legame fosfodiesterico con il DNA cellulare (accettore). Quindi essi non inibiscono il 3'-processing e non competono per il legame con il DNA virale. Al contrario, i di-chetoacidi bifunzionali potrebbero legare l'enzima sia nel sito di catalisi del 3'-processing (sito donatore) che in quello dello strand transfer⁷⁶.

A partire dalla serie di composti poliossidrilati, analoghi della curcumina, visti in precedenza risultati molto attivi sia nei saggi enzimatici che cellulari, però molto citotossici, sono stati quindi progettati e sintetizzati una serie di composti correlati ai derivati cinnamoilici allo scopo di ottenere nuovi inibitori dell'integrasi, che mostrassero una maggiore attività inibitoria nei confronti della replicazione dell'HIV-1 e che fossero dotati allo stesso tempo di minore citotossicità. Si è arrivati così alla sintesi degli acidi 6-(1-arylmetil-1*H*-pirrol-2-il)-2,4-dioxo-5-esenoici (**29**):



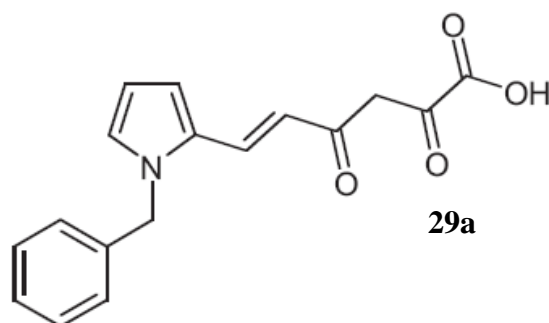
Questi composti presentano una funzione cinnamoilica in cui l'anello benzenico è sostituito con una diversa porzione aromatica, in particolare con un

fenilmetil pirrolo, presente anche nel derivato Merck L-731,988. Inoltre conservano l'1,3-dicheto gruppo e la funzione carbossilica acida, fondamentali per ottenere composti ad elevata attività anti-HIV. L'inserimento di un gruppo etilenico tra il fenilmetil pirrolo e la porzione 2,4-dioxobutanoica, rende questi composti viniloghi del derivato Merck L-731,988.

Gli acidi arildiooesenoici, composto 29, derivati dalla parziale sovrapposizione della porzione cinnamoilica e dioxobutanoica, possono quindi essere considerati ibridi di queste due diverse classi di inibitori dell'integrasi, i cinnamoil derivati e i dichetoacidi. Dall'analisi strutturale dei composti poliossidrilati e degli acidi arildichetooesenoici, si può notare che questi inibitori dell'integrasi presentano delle caratteristiche comuni, necessarie ai fini dell'attività inibitoria:

1. il gruppo cinnamoilico, presente sia nei prodotti naturali che in quelli sintetici;
2. 1,3-dicheto gruppo, presente nella serie degli arildichetoacidi;
3. la funzione carbossilica, riscontrata nei prodotti naturali (acido L-cicorico), nei dichetoacidi (L-731,988) e nei derivati a struttura Cinnamoilica (cp 17-18);
4. la porzione aromatica.

Gli acidi arildichetoeseoici **29** presentano tutte queste caratteristiche strutturali, e si sono rivelati dei potenti agenti anti-HIV⁶⁷. Inizialmente è stato sintetizzato il composto non sostituito sull'anello benzenico (**29a**) e ne è stata determinata l'attività enzimatica e cellulare.



29a ha mostrato un'ottima azione inibente nei confronti dell'integrasi, sia sul 3'-processing, con una IC_{50} di 7.9 μM , che sullo strand transfer con una IC_{50} pari a 7 μM ed una attività antivirale, con una EC_{50} di 1.5 μM paragonabile a quella del composto Merck L-731,988, che presenta una EC_{50} di 1 μM . Dato l'ottimo risultato ottenuto con questo composto, per studiare le relazioni struttura ed attività, è stata sintetizzata una serie di derivati che hanno diversi sostituenti in posizione 4 dell'anello benzenico. È stata scelta proprio la

posizione 4 dell'anello benzenico per analogia al derivato Merck L-731,988. L'introduzione di gruppi lipofili come il metile, e il cloro o di gruppi a diverso effetto elettronico come il nitro o il metossile hanno portato al completo annullamento dell'attività, sia in saggi enzimatici che cellulari. L'unico che ha mantenuto una discreta attività su cellule infettate è risultato il derivato 4-F sostituito ($EC_{50} = 11 \mu M$).

Ciò rende evidente che fattori come la lipofilia, il rilascio e l'attrazione elettronica dei vari gruppi non sono determinanti per l'attività biologica. Si potrebbe quindi ipotizzare che un ruolo fondamentale per l'attività biologica sia giocato dall'ingombro sterico. Infatti i migliori risultati sono stati ottenuti attraverso la sostituzione dell'idrogeno in 4 con un fluoro, il più piccolo fra tutti i sostituenti scelti⁶⁷.

Per approfondire gli studi in questo interessante campo di indagine, è stato realizzato un approccio preliminare di molecular modeling, studiando l'interazione dei nuovi derivati con il sito attivo dell'integrasi dell'HIV-1, scegliendo il derivato non sostituito sull'anello benzenico (**29a**) come composto rappresentativo della serie degli acidi aril-diooesenoici. Il composto Merck L-731,988 (giallo) e l'acido arildiooesenoico di riferimento **29a** (viola) sono stati allineati sovrapponendo atomo per atomo le loro porzioni dichetoacide con quella del 5-CITEP (arancio). Si può notare che fra i composti così sovrapposti si

verifica un buon grado di allineamento. Inserendo il derivato Merck e l'acido arildiooesenoico nel sito attivo dell'integrasi, nel quale è cocristallizzato il 5-CITEP si nota che la parte farmacofora di questi inibitori ha una interazione ligando-recettore comune, caratteristica di tutti e tre i composti.

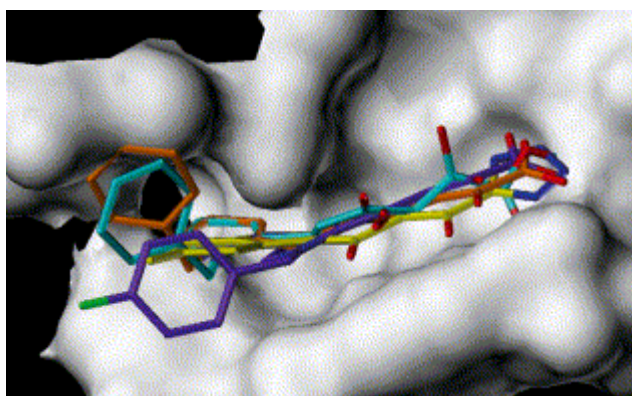
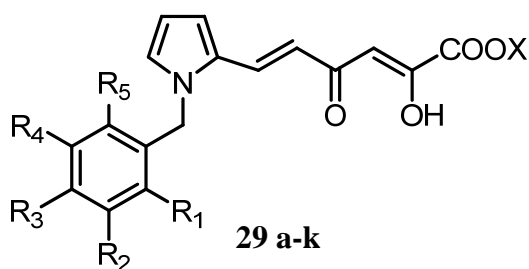


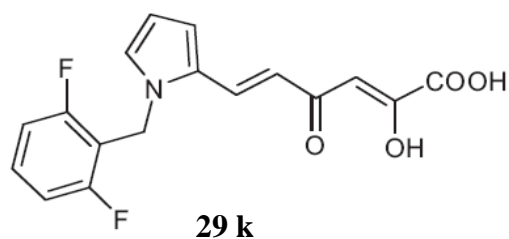
Figura 19. Sovrapposizione i L-731,988 (giallo), acido arildiooesenoico **29a** (viola) e 5-citep (arancio) nel sito catalitico dell'integrasi

Un aspetto di notevole interesse è la differente disposizione della catena laterale 4-fluorobenzilica di L-731,988 e della porzione 5-cloroindolica del 5-CITEP rivolte verso l'esterno della cavità recettoriale, rispetto al gruppo benzilico dell'acido arildiooesenoico, che invece rimane intrappolato nella tasca dell'integrasi. Dall'immagine rappresentata in **figura 19**, si può razionalizzare la scarsa attività dei composti 4-sostituiti. Si può inoltre ipotizzare che l'introduzione di un sostituito in posizione 3 o in posizione 2 dell'anello benzenico, possa favorire l'interazione con la tasca del sito attivo dell'integrasi,

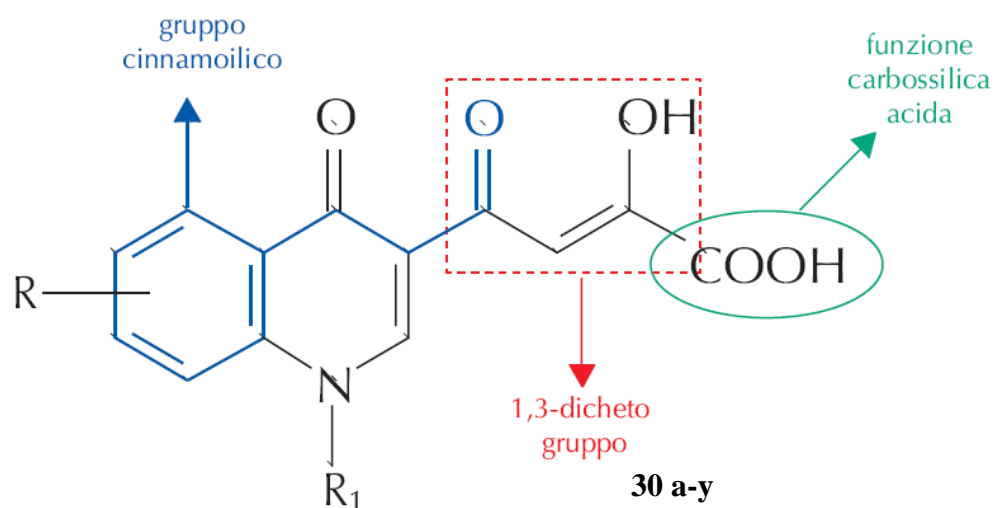
aumentando l'attività inibitoria nei confronti dell'enzima stesso. Inoltre è importante indagare gli effetti sull'attività biologica di una doppia sostituzione, in posizione 3,5 o ancora in posizione 2,6, che genera degli analoghi conformazionalmente bloccati. Si è pertanto giunti alla progettazione di acidi arildiooesenoici (**29a-m**) al fine di valutare a fondo l'influenza di queste sostituzioni sull'attività biologica di tali derivati.



Nell'ambito di questa serie di composti i risultati migliori sono stati ottenuti con il derivato 2,6-difluoro sostituito **29k**, che presenta una EC_{50} pari a $0.35 \mu\text{M}$, una CC_{50} pari a $37\mu\text{M}$, ed un indice di selettività S.I.= 106, dimostrandosi tre volte più attivo del derivato Merck **L-731,988** preso come riferimento.



L'elevato valore di attività riscontrato per il derivato **29k** potrebbe essere dovuto al fatto che la doppia sostituzione in posizione 2,6 sull'anello benzilico genera un composto conformazionalmente bloccato, che probabilmente si lega più efficacemente al sito attivo enzimatico e ne potenzia l'attività. Sulla base di questa ipotesi strutturale, ci si è rivolti alla progettazione di nuovi potenziali inibitori dell'integrasi dell'HIV-1, conservando inalterate le caratteristiche strutturali necessarie all'attività inibitoria, e cioè la funzione cinnamoilica, l'1,3-dichetogruppo e la funzione carbossilica acida. Inoltre si è pensato di incorporare la porzione cinnamoilica in un sistema rigido, ed in particolare in un anello oxochinolinico, in modo da generare analoghi conformazionalmente bloccati degli acidi arildiooesenoici **29**, per migliorare l'interazione con il sito catalitico enzimatico e l'attività biologica. La scelta dell'anello oxochinolinico è stata indotta dall'osservazione che numerosi farmaci di uso clinico sono derivati della chinolina, in particolare i chinoloni antibatterici.



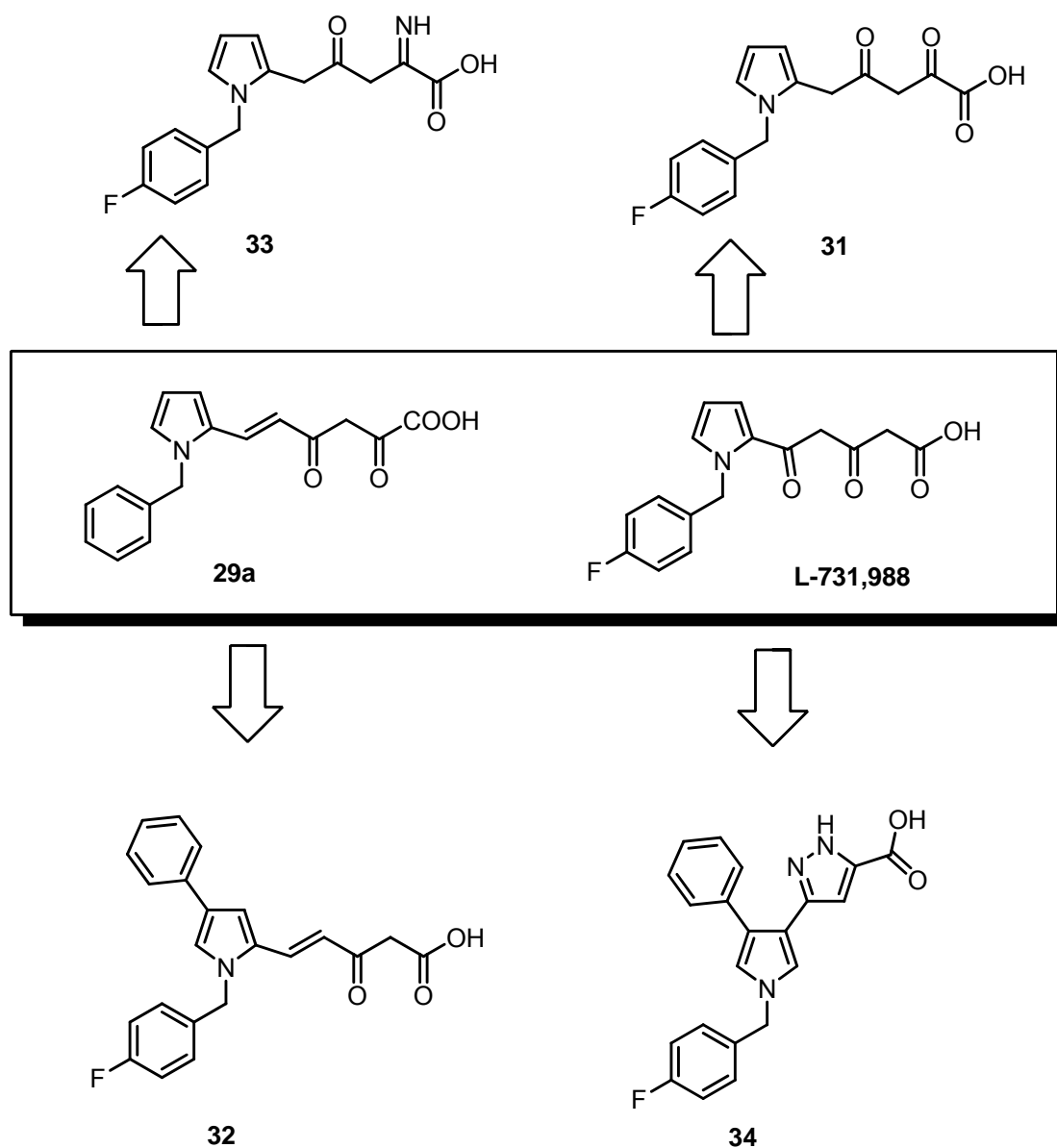
Si è giunti così alla progettazione degli acidi chinolinildioxobutanoici **30 a-y** variamente sostituiti sull'atomo di azoto e sulle posizioni 6 e 7 dell'anello oxochinolinico. I composti così progettati, pur conservando inalterata la porzione farmacoforica dichetoacida, hanno permesso, allo stesso tempo, di incorporare la funzione cinnamoilica in un sistema rigido, conformazionalmente bloccato, quale la struttura oxochinolinica, in modo da favorire l'interazione selettiva con il sito attivo enzimatico. I buoni risultati ottenuti con i derivati di **29** e **30**, ci hanno condotto ad ulteriori indagini sugli arildichetoacidi come inibitori dell'integrasi. In particolare, dato che i ricercatori della Merck hanno messo in evidenza che l'orientamento ottimale tra la catena di chetoacida ed il gruppo benzilico dei dichetoacidi è di circa 120°, è stato progettato il derivato 4-[4-fenil-1-(4-fluorofenil)metil-1*H*-pirrol-3-il]-2,4-dioxobutanoico (**32**).

In effetti questo composto mostra le seguenti caratteristiche chimiche:

- 1) un fenile in posizione 4 del pirrolo;
- 2) un gruppo benzilico in posizione 1 dell'anello pirrolico;
- 3) una catena dichetoacida disposta in posizione 1,3 rispetto alla porzione benzilica.

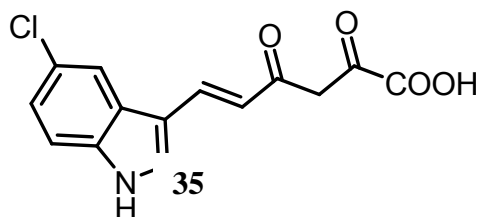
Successivamente sul derivato **31** sono state progettate le seguenti modifiche:

- 4) allungamento della catena dioxobutanoica, che porta all'acido 4-[4-fenil-1-(4-fluorofenil)metil-1*H*-pirrol-3-il]-2,4-dioxo-5-esenoico, composto **32**;
- 5) sostituzione della catena dioxobutanoica con un gruppo chetoimminico (acido 2-ammino-4-[1-[(4-fluorofenil)metil]-4-fenil-1*H*pirrol-3-il]-4-oxo-2-butenico, composto **33**;
- 6) introduzione di un pirazolo allo scopo di mascherare il dicheto gruppo (acido 5-[1-[(4-fluorofenil)metil]-4-fenil-1*H*-pirrol-3-il]-1*H*-pirazol-3-carbossilico, composto **34**;



Come ulteriore sviluppo della ricerca da acidi eteroaril dichetoesenici come agenti anti-integrasi, è stata pianificata la sintesi dell'acido 6-(5-cloro-1*H*-indol-3-il)-2,4-dioxo-5-esenico, composto **35**. Questo composto è stato progettato

come risultato della fusione della porzione 5-cloroindolica del 5-CITEP con la catena dichetoeseenoica degli acidi arildichetoeseenoici.

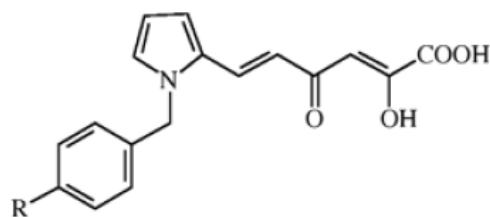


Tutti i derivati progettati, composti **29-35**, ed i corrispondenti esteri etilici sono stati saggiati in test enzimatici per determinarne l'inibizione nel 3'-processing e nello strand transfer, in saggi su cellule infettate da HIV-1, e su cellule non infettate, per determinare la citotossicità, usando come sostanza di riferimento il derivato Merck **L-731,988**. Inizialmente sono stati sintetizzati i composti a struttura pirrolil dichetoeseenoica **29**. Essi risultano ibridi tra due classi di inibitori dell'integrasi riportati in letteratura: gli arildichetoacidi ed i cinnamoil-derivati. Essi infatti presentano la funzione cinnamoilica modificata a livello della porzione aromatica dall'introduzione del fenilpirrolo e la catena dioxobutanoica parzialmente sovrapposta alla porzione cinnamoilica. Il risultato di questa parziale sovrapposizione è l'elongazione della catena dioxobutanoica

grazie all'inserimento di un ponte etilenico tra il fenilpirrolo e il carbonile del dichetoacido: si crea così una catena 2,4-dioxo-5-esenoica. Per valutare la efficacia di tale classe di inibitori si è pensato di sintetizzare all'inizio il composto non sostituito sull'anello benzenico (**29a**) e di determinarne l'attività enzimatica e cellulare. Esso ha mostrato un'ottima azione inibente nei confronti dell'integrasi, sia sul 3'-processing, con una IC_{50} di 7.9 μM , che sullo strand transfer con una IC_{50} pari a 7 μM ed una attività antivirale, con una EC_{50} di 1.5 μM paragonabile a quella del composto Merck **L-731,988**, che presenta una EC_{50} di 1 μM . Dato l'ottimo risultato ottenuto con questo composto, per studiare le relazioni tra struttura ed attività biologica, è stata sintetizzata una serie di derivati che hanno diversi sostituenti in posizione 4 dell'anello benzenico. È stata scelta proprio la posizione 4 dell'anello benzenico per analogia al derivato Merck.

L-731,988, che presenta un fluoro in tale posizione. L'introduzione di gruppi lipofili come il metile (**29p**), di gruppi ingombranti come il cloro (**29o**) o di gruppi a diverso effetto elettronico come il nitro (**29r**) o il metossile (**29q**) hanno portato al completo annullamento dell'attività, sia in saggi enzimatici che cellulari. L'unico che ha mantenuto una discreta attività su cellule infettate è risultato il derivato 4-F sostituito (**29n**) (EC_{50} = 11 μM) (**tabella 2**).

Tabella 2 Attività su cellule infettate e citotossicità dei derivati **29n-r**



Cp	R	CC ₅₀ (μM)	EC ₅₀ (μM)	SI	IC ₅₀ (μM)	
					3'-P	ST
29n	F	80	11	7.3	> 10	> 10
29o	Cl	95	>95	-	> 10	> 10
29p	CH ₃	41	>41	-	> 10	> 10
29q	CH ₃ O	78	>78	-	> 10	> 10
29r	NO ₂	33	>33	-	> 10	> 10

Ciò rende evidente che fattori come la lipofilia, il rilascio e l'attrazione elettronica dei vari gruppi non sono determinanti per l'attività biologica. Si potrebbe quindi ipotizzare che un ruolo fondamentale per l'attività biologica sia giocato dall'ingombro sterico.

Infatti i migliori risultati sono stati ottenuti attraverso la sostituzione dell'idrogeno con un fluoro, il più piccolo fra i sostituenti scelti⁶⁷.

Dati i risultati ottenuti con i derivati arilmetil-2,4-dioxo-5-esenoici sostituiti in posizione 4 dell'anello aromatico, sono stati sintetizzati derivati 2-sostituiti, 3-sostituiti, 2,6-disostituiti, e 3,5-disostituiti. Questi composti hanno evidenziato in generale una buona attività inibitoria nei confronti del target biologico in questione. Gli acidi 2-sostituiti (**29b-f**) e i corrispondenti esteri etilici non presentano sostanziali differenze di attività contro cellule infettate da HIV-1.

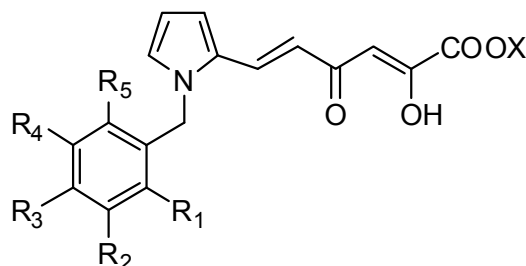
Visto che ai fini dell'attività mostrata su cellule infettate non sembrava giocare un ruolo rilevante la presenza della porzione carbossilica in luogo di quella esterea e viceversa, a questo punto è risultato logico pensare come la presenza di diversi sostituenti sul benzile sia la sola discriminante al fine di ottenere composti con diversa attività in questa serie omologa. Per descrivere le SAR di questa serie di composti si è preso in considerazione il composto **29a** come lead compound. Le relazioni struttura-attività in questa serie di dichetoacidi sono descritte in base all'attività antivirale poichè non sono ancora disponibili i dati relativi all'inibizione enzimatica. Ciò nonostante si può asserire con ragionevole certezza che tali composti inibiscono l'integrasi. Infatti il nostro lead compound **29a** è risultato un inibitore dell'integrasi in test enzimatici. Ai

fini dell'attività sembra giocare un ruolo importante la presenza di un sostituito ad attrazione elettronica in posizione 2, come il fluoro nei composti **29b** ($EC_{50}=1\mu M$) anche se l'effetto di attrazione elettronica non appare evidente come unica prerogativa intesa a giustificare l'attività dei composti in questione. Infatti nel composto **29c** ($EC_{50}=4.2\mu M$) la presenza del cloro, con le medesime caratteristiche elettroniche del fluoro, non porta ad un aumento dell'attività, bensì a composti meno attivi, forse per l'effetto ingombrante del medesimo. Bisogna comunque sottolineare che nel passaggio dal fluoro al cloro si ha un aumento della citotossicità e una conseguente diminuzione della selettività dei derivati **29c**. Anche la presenza del metile in posizione 2 dell'anello benzenico, che presenta al contrario un effetto di rilascio elettronico, ha portato a derivati con attività paragonabile a quella di **29b** e con il fluoro in posizione 2. Infatti il derivato **29d** presenta EC_{50} pari a $0.7\mu M$ e molto probabilmente la sua attività non dipende dagli effetti elettronici dei sostituenti sull'anello. Il derivato **29d** presenta un S.I.= 50, con una buona selettività nei confronti delle cellule infettate. La presenza di sostituenti a rilascio elettronico con un certo grado di lipofilia quali il gruppo OEt e OCH₃ nei composti **29e**, **29f**, ($EC_{50}=2\mu M$, $EC_{50}=2.5\mu M$, rispettivamente), non ha portato a composti più attivi di quelli già menzionati. Successivamente sono stati sintetizzati gli acidi 3-sostituiti

29g-m. I valori di EC_{50} evidenziano come la sostituzione in posizione 3 dell'anello aromatico fornisca composti con minore attività rispetto agli acidi **29b-f**. Anche in questo caso non esistono grandi differenze di attività tra gli acidi e gli esteri. Infatti per ogni coppia di derivati, i valori di EC_{50} sono molto simili tranne che per il derivato **29j** che mostra una $EC_{50} = 6.5 \mu M$. In questo gruppo di composti non sembrano giocare un ruolo chiave le differenze steriche ed elettroniche dei sostituenti. Nei derivati 3-sostituiti non sembra fondamentale ai fini dell'attività la presenza del fluoro in luogo del cloro, visto che i valori di EC_{50} si mantengono molto simili (**29g**: $EC_{50} = 7 \mu M$, , **29h**: $EC_{50} = 7 \mu M$). Persino la presenza di un gruppo metile (**29i**) o di un gruppo metossile (**29j**) al posto del fluoro o del cloro non cambia di molto l'attività di tali composti. Tra i derivati 3-sostituiti quello che presenta maggiore attività è il composto **29j** con una $EC_{50} = 6.5 \mu M$ in cui compare il gruppo metossile a rilascio elettronico. Gli altri derivati 3-sostituiti presentano attività molto simili e valori di S.I. abbastanza bassi. Tra i composti disostituiti mostra una ottima attività il derivati 2,6-difluoro sostituito **29k**. I derivati 3,5-disostituiti **29l,m** risultano invece inattivi in saggi cellulari, ad eccezione del composto **29l** che mostra una debole attività, con una $EC_{50} = 16 \mu M$. La presenza di due atomi di fluoro in posizione 2,6 quindi fornisce i composti più attivi (**29k**: $EC_{50} = 0.35 \mu M$). Anche in questo caso è

importante sottolineare come la presenza di un gruppo ad attrazione elettronica poco ingombrante come il fluoro sia fondamentale per l'attività rispetto a gruppi più ingombranti come il cloro. Confrontando l'attività dei composti sintetizzati con il derivato **L-731,988** sintetizzato dalla Merck Company, risulta evidente come uno dei composti da noi sintetizzati, ed in particolare il derivato 2,6-difluorosostituito (**29k**) sia ben tre volte più attivo del composto di riferimento. Questo testimonia che ai fini dell'attività risulta fondamentale la presenza del fluoro in posizione *orto* dell'anello aromatico. I buoni risultati ottenuti con il derivato 2-fluoro sostituito **29b** sono notevolmente incrementati dall'introduzione di un secondo atomo di fluoro in posizione 6 (**tabella 3**).

Tabella 3. Attività su cellule infettate e citotossicità dei derivati 29 a-m



Cp	R ₁	R ₂	R ₃	R ₄	R ₅	X	CC ₅₀	EC ₅₀	S.I.
29a	H	H	H	H	H	H	61	1.5	7.0
29b	F	H	H	H	H	H	49	1.0	49
29c	Cl	H	H	H	H	H	52	4.2	12
29d	Me	H	H	H	H	H	35	0.7	50
29e	OMe	H	H	H	H	H	56	2	28
29f	OEt	H	H	H	H	H	32.5	2.4	135
29g	H	F	H	H	H	H	57.0	7.0	8.0
29h	H	Cl	H	H	H	H	54	7	7.7
29i	H	Me	H	H	H	H	46	7.4	6
29l	H	Cl	H	Cl	H	H	45.6	16	2.8
29m	H	Me	H	Me	H	H	47	33	-
L-731,988							54	1,5	36

L'elevato valore di attività riscontrato per il derivato **29k** potrebbe essere dovuto al fatto che la doppia sostituzione in posizione 2,6 sull'anello benzilico genera un composto conformazionalmente bloccato, che probabilmente si lega più efficacemente al sito attivo enzimatico e ne potenzia l'attività. I buoni

risultati ottenuti con i derivati **29**, ci hanno condotto ad ulteriori indagini sugli arildichetoacidi come inibitori dell'integrasi.

In particolare, dato che i ricercatori della Merck hanno messo in evidenza che l'orientamento ottimale tra la catena di chetoacida ed il gruppo benzilico dei dichetoacidi è di circa 120°, è stato sintetizzato il derivato 4-[4-fenil-1-(4-fluorofenil)metil-1*H*-pirrol-3-il]-2,4-dioxobutanoico (**31**). In effetti questo composto mostra le seguenti caratteristiche chimiche:

- 1) un fenile in posizione 4 del pirrolo;
- 2) un gruppo benzilico in posizione 1 dell'anello pirrolico;
- 3) una catena dichetoacida disposta in posizione 1,3 rispetto alla porzione benzilica.

Esso è risultato attivo sia in saggi enzimatici ($IC_{50} = 0.026 \mu M$) che cellulari ($EC_{50} = 9 \mu M$, S.I. >22). Successivamente sul derivato **31** sono state effettuate le seguenti modifiche:

- 1) allungamento della catena dioxobutanoica, che porta all'acido 4-[4-fenil-1-(4-fluorofenil)metil-1*H*-pirrol-3-il]-2,4-dioxo-5-esenoico (**32**);

- 2) sostituzione della catena dioxobutanoica con un gruppo chetoimminico (acido 2-ammino-4-[1-[(4-fluorofenil)metil]-4-fenil-1*H*pirrol-3-il]-4-oxo-2-butenico, (**33**);
- 3) introduzione di un pirazolo allo scopo di mascherare il dicheto gruppo (acido 5-[1-[(4-fluorofenil)metil]-4-fenil-1*H*-pirrol-3-il]-1*H*-pirazol-3-carbossilico, (**34**).

Tali composti hanno mostrato una discreta attività biologica, comunque inferiore a quella dei derivati **29** e **30**, ad eccezione del derivato pirazolico, che è risultato completamente inattivo (**tabella 5**). Probabilmente l'introduzione di un sistema rigido, che mima la funzione cheto-enolica, non è adeguata al mantenimento dell'attività nei confronti dell'integrasi.

Nell'ambito della serie dei derivati pirrolici, comunque, i migliori risultati sono stati ottenuti con gli acidi pirrolidichetoenoici **29**, ed in particolare con il derivato 2,6-difluorosostituito **29k**. L'elevato valore di attività riscontrato per il derivato **29k** potrebbe essere dovuto al fatto che la doppia sostituzione in posizione 2,6 sull'anello benzilico genera un composto conformazionalmente bloccato, che probabilmente si lega più efficacemente al sito attivo enzimatico e ne potenzia l'attività.

Parte III

Progettazione di inibitori dichetoacidi bifunzionali (BDK)

Sulla base di questa ipotesi strutturale, successivamente si è pensato di progettare una nuova serie di potenziali inibitori dell'integrasi di HIV-1 che conservasse inalterate le caratteristiche strutturali necessarie all'attività inibitoria, e cioè la funzione cinnamoilica, l'1,3-dichetogruppo e la funzione carbossilica acida. Pertanto si è pensato di incorporare la porzione cinnamoilica in un sistema rigido, ed in particolare in un anello oxochinolinico, in modo da generare analoghi conformazionalmente bloccati degli acidi arildiooesenoici **29**, per migliorare l'interazione con il sito catalitico enzimatico e l'attività biologica. Da un'analisi iniziale dei valori ottenuti, si può notare che, in generale, tutti i derivati acidi chinolonici (**29a-o**) mostrano un profilo di attività e una capacità inibitoria nei confronti dell'integrasi migliore rispetto ai corrispondenti derivati esterei. Questo dato sperimentale conferma ulteriormente l'importanza della porzione dichetoacida come caratteristica strutturale fondamentale ai fini

dell'attività biologica. Si è pensato perciò di valutare l'influenza delle modifiche strutturali del sostituente in posizione N-1 dell'anello oxochinolinico sulla capacità inibitoria⁷⁵.

Inizialmente è stato sintetizzato il composto **30a**, introducendo un gruppo benzilico non sostituito, e ne è stata determinata l'attività enzimatica e cellulare. Il composto ha mostrato un'ottima azione inibente nei confronti dell'integrasi in test enzimatici, sia sul 3'-processing ($IC_{50} = 4.0 \mu M$), che sullo strand transfer ($IC_{50} = 0.034 \mu M$). Presenta inoltre una notevole attività inibitoria nei confronti delle cellule infettate con una $EC_{50} = 9.6 \mu M$, ed un indice di selettività > 20.8 . È necessario sottolineare che passando dalla serie pirrolica a quella chinolonica si è avuto un notevole incremento dell'attività, stimabile intorno ai 2-3 ordini di grandezza (**29a** $IC_{50} ST = 7.9 \mu M$, **30a**: $IC_{50} ST = 0.034 \mu M$).

Dato l'ottimo risultato ottenuto, sono stati sintetizzati una serie di derivati acidi con diversi sostituenti sull'anello benzilico in N-1 (**30b-k**), al fine di valutare l'influenza che queste sostituzioni hanno sull'attività biologica. L'introduzione di un atomo di piccole dimensioni e ad attrazione elettronica come il fluoro in posizione 4 ha portato ad ottimi risultati: il 4-fluoro derivato **30d**, risulta infatti molto attivo nell'inibire lo strand transfer ($IC_{50} = 0.0313 \mu M$), con attività paragonabile al derivato non sostituito **29a**.

La sostituzione in posizione 4 con un gruppo ingombrante come il cloro, ha portato invece ad un composto **30e** moderatamente attivo, con valori di IC₅₀ pari a 22 μM (3'-P) e 0.54 μM (ST). L'introduzione in posizione para di un gruppo ad elevato ingombro sterico e a rilascio elettronico come il metossile (**29i**) o di un gruppo elettron-attrattore come il nitro (**30j**), non provoca un sostanziale incremento di attività. Si ottengono comunque composti moderatamente attivi nei saggi enzimatici, con una IC₅₀ pari rispettivamente a 28 e 27 μM per il 3'-processing e 0.51 e 0.64 μM per lo strand transfer, paragonabili a quella del derivato 4-cloro sostituito **30e**. Ciò rende evidente che fattori come la lipofilia, il rilascio o l'attrazione elettronica dei vari sostituenti esaminati in questa posizione non influenzano significativamente l'attività biologica. Si può solo notare che le attività dei composti che presentano un qualunque sostituente nella posizione 4 sono inferiori a quella del composto non sostituito **30a**. Si potrebbe quindi ipotizzare che un fattore determinante per l'attività biologica sia l'ingombro sterico: i migliori risultati sono stati infatti ottenuti attraverso la sostituzione con il fluoro, che risulta essere l'atomo più piccolo fra tutti quelli scelti per la sostituzione in posizione 4 dell'anello benzenico. Lo spostamento del fluoro dalla posizione 4 alla posizione 2 dell'anello benzenico ha portato ad un incremento dell'attività: il composto così

ottenuto **30b** presenta infatti valori di IC_{50} pari a 1.2 μM per il 3'-processing e 0.016 μM per lo strand transfer, un valore di EC_{50} nei saggi cellulari uguale a 9.04 μM , con indice di selettività pari a 17.8. Infine lo spostamento del fluoro dalla posizione 2 alla posizione 3, porta ad un composto ugualmente molto attivo **30c**, con un profilo di attività enzimatica paragonabile a quella del composto **30b** ($IC_{50} = 4.6 \mu M$ per il 3'-P e 0.015 μM per lo ST), ed una EC_{50} pari 37.8 μM . I composti **30b** e **30c** risultano infatti i più attivi della serie tra gli acidi chinolinil-2,4-dioxobutanoici da noi sintetizzati, con valori di attività nei confronti dello strand transfer circa doppi rispetto al derivato non sostituito **30a**. Inoltre, come si può facilmente notare dai valori di EC_{50} nei saggi cellulari e dai rispettivi indici di selettività, i composti **30a** e **30b** presentano una notevole attività inibitoria nei confronti della replicazione virale in cellule infettate da HIV-1, a concentrazioni di inibitore notevolmente inferiori a quelle citotossiche.

Buoni risultati sono stati ottenuti anche attraverso la doppia sostituzione con il fluoro in posizione 2,6: il composto **30k** risulta infatti molto attivo nei saggi enzimatici con un valore di IC_{50} di 0.7 μM per il 3'-processing e di 0.019 μM per lo strand transfer.

Mentre le modifiche a livello dell'anello benzilico hanno portato ad un miglioramento dei valori di IC_{50} , diversa è la situazione per quanto riguarda

l'effetto della sostituzione del nucleo benzilico sull'azoto della oxochinolina con altri gruppi. Introducendo infatti in questa posizione un gruppo fenilbutilico, e quindi allungando la catena alchilica tra il nucleo oxochinolinico e quello aromatico, si ottiene il composto **30l** dotato di buona attività enzimatica ma inattivo nei test su cellule infettate ($IC_{50} > 333 \mu M$ per il 3'P e $< 0.45 \mu M$ per lo ST, $CC_{50} = 180 \mu M$ e $EC_{50} > 180 \mu M$). Analoghi risultati sono stati ottenuti inserendo un anello aromatico più ingombrante del fenile come il gruppo naftilico (**30m**), od anelli eterociclici come la piridina (**30n**) e la chinolina (**30o**). Infatti, nei saggi enzimatici, tali derivati risultano tutti meno attivi dei corrispondenti derivati benzilici. Si può quindi ipotizzare che la presenza di un eterociclo nel sostituente in posizione 1 non influisca positivamente sull'attività di questi composti. Inoltre l'elevato ingombro sterico di alcuni gruppi (**30l**, **30m** e **30o**) probabilmente non permette l'instaurarsi di interazioni favorevoli con il sito attivo dell'integrasi. Infine, dall'analisi dei valori di EC_{50} e CC_{50} , si può facilmente osservare che, per i composti **30l-o**, l'attività inibitoria nei confronti della replicazione virale in cellule infettate, si manifesta solo a concentrazioni che vicine a quelle citotossiche. È molto importante sottolineare che i derivati sintetizzati, analogamente agli altri dichetoacidi riportati in letteratura, sono attivi selettivamente nei confronti dello strand transfer. Infatti i derivati più attivi

30a-d sintetizzati mostrano un rapporto IC_{50} 3'-P/ IC_{50} ST molto elevato (> 70 per i derivati **30b** e **d**, >100 per **30a** e >300 per **30c**).

Questo fattore ci fa presupporre che interagiscano specificatamente con il sito di catalisi dello strand transfer, inibendo solamente tale step del processo di integrazione. Riguardo le sostituzioni sulle posizioni 6, 7 e 8 dell'anello oxochinolinico, si può osservare che, in generale, si ottengono derivati ad attività paragonabile rispetto ai composti non sostituiti e che gli esteri risultano meno attivi degli acidi carbossilici corrispondenti (nei test enzimatici). L'introduzione di un atomo di fluoro in posizione 6 ha portato al derivato **30p** che ha mostrato una buona attività nei confronti dello strand transfer ($IC_{50} = 0.047 \mu M$), pur risultando meno attivo dell'analogo non sostituito **30d** ($IC_{50} = 0.0313 \mu M$). L'inserimento del cloro in posizione 6 conduce invece ad un composto **30q** poco attivo nei confronti dello ST ($IC_{50} = 0.124 \mu M$). Evidentemente l'introduzione di un atomo di grosse dimensioni come il cloro in posizione 6 non porta ad interazioni favorevoli con l'enzima. La sostituzione in posizione 7 con il fluoro (**30s**) o con il cloro (**30u**) ha portato a derivati abbastanza attivi nei confronti dell'integrasi con valori di IC_{50} paragonabili o addirittura superiori, come nel caso del derivato 7-cloro sostituito ($IC_{50} = 0.0210 \mu M$), rispetto al composto non sostituito **30d**. La doppia sostituzione con un atomo di cloro in posizione 6,7 ha

portato ad un peggioramento dell'attività: il derivato **30z** ($IC_{50} = 0.032 \mu M$) risulta infatti più attivo del composto 6-cloro sostituito **30p** ($IC_{50} = 0.124 \mu M$), ma risulta meno attivo dell'isomero in posizione 7 (**30s**), che presenta una IC_{50} sullo ST pari a $0.0210 \mu M$. Ottimi risultati sono stati invece ottenuti con l'introduzione di atomi di fluoro o di cloro in posizione 8 dell'anello oxochinolinico. Il derivato **30w**, 8-fluoro sostituito risulta infatti il più attivo della serie fluorurata con una IC_{50} pari a $0.025 \mu M$, risultando più attivo anche del composto non sostituito **30d**. I migliori risultati sono stati comunque ottenuti con il derivato 8- cloro sostituito **30x**. Esso presenta una IC_{50} per lo ST pari a $0.014 \mu M$, doppia rispetto a quella del derivato non sostituito **30d** e paragonabile a quella dei derivati **30b** e **30c**. L'introduzione di un anello pirrolidinico in posizione 7 ha portato ad una diminuzione dell'attività inibitoria nei confronti dell'integrasi: si hanno infatti valori di IC_{50} pari a $0.079 \mu M$, circa 6 volte superiori rispetto a quelli del derivato 8-cloro sostituito. È comunque importante sottolineare che tale derivato risulta attivo nei saggi cellulari, e questo offre l'opportunità di indagare a fondo sulla possibilità di effettuare altre sostituzioni con ammine, in modo da valutare l'effettivo impiego terapeutico di tali composti. Ottimi risultati sono stati ottenuti con il derivato **30r**. Tale composto mostra due catene dioxobutanoiche in posizione 3,6 dell'anello

oxochinolinico, quindi può essere considerato un di chetoacido bifunzionale. Esso è risultato il più attivo della serie degli acidi 4- (1,4-diidro-1-aril-4-oxochinolin-3-il)-2,4-dioxobutanoici (**30**), mostrando un'ottima attività nei confronti dell'integrasi con valori di $IC_{50} = 0.012 \mu M$ e risultando dalle 17 alle 540 volte più attivo di alcuni dichetoacidi bifunzionali riportati in letteratura. Esso inoltre è attivo anche nei confronti delle cellule infettate da HIV-1 ($EC_{50} = 4.29 \mu M$), risulta scarsamente citotossico ($CC_{50} > 200 \mu M$) e mostra un indice di selettività >47 : può essere quindi considerato un ottimo derivato sul quale investire per una potenziale applicazione terapeutica. (**tabella 6**).

L'elevata attività di tale composto potrebbe essere legata al fatto che esso presenta l'anello aromatico centrale 1,3 disostituito in modo da formare un angolo di 120° tra il sostituente benzilico e il gruppo dichetoacido in 3. Tali requisiti strutturali sembrano aumentare molto l'attività inibitoria dei derivati arildioxobutanoici nei confronti dell'integrasi.

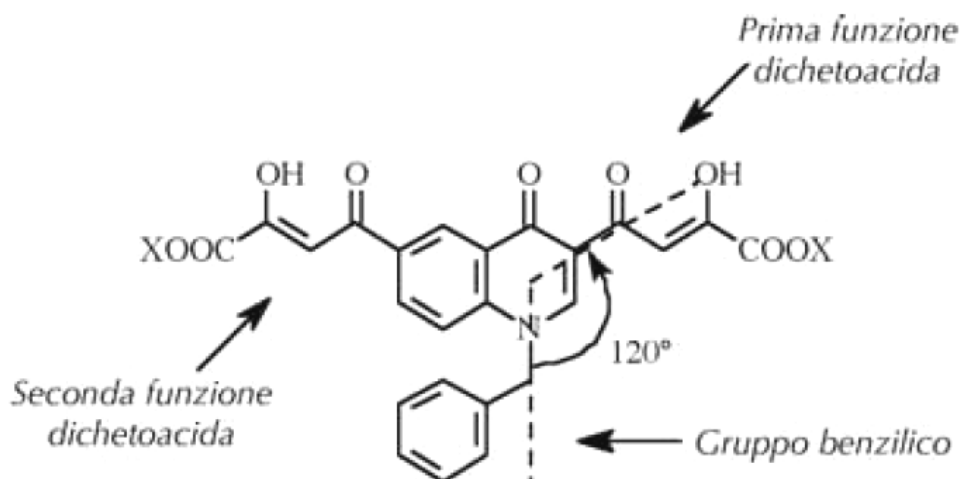


Figura 20. Requisiti strutturali dei derivati a struttura chinolin di-chetoacida

Il composto **30r** presenta il gruppo benzilico sulla posizione 1 dell'anello oxochinolinico che consente di ottenere un derivato 1,3 disostituito che ben si adatta ai requisiti geometrici per un'elevata attività anti-integrasi. Una seconda funzione dioxobutanoica inoltre può essere facilmente introdotta attraverso un acetil derivato intermedio. Come è mostrato in **figura 23**, i nuovi dichetoacidi bifunzionali presentano un'elevata attività sia nei confronti dello strand transfer che del 3'-processing. In particolare il derivato **30r** risulta il più potente della serie dei dichetoacidi bifunzionali, con valori di IC_{50} per lo ST di circa 15 nM, inoltre risulta selettivo nei confronti dello strand transfer con valori di IC_{50} per lo ST 20 volte più bassi rispetto a quelli del 3'-P.

È molto importante sottolineare che il composto **30r** inibisce l'integrasi con potenza simile sia in presenza di Mg^{2+} che in presenza di Mn^{2+} . L'attività in presenza di Mg^{2+} è molto importante in quanto tutti i derivati attivi solamente in presenza di Mn^{2+} sono generalmente citotossici ed inattivi in saggi cellulari.⁷⁴ Inoltre la selettività nei confronti dello ST, aggiunta all'attività inibente in presenza di Mg^{2+} , e ai valori di IC_{50} sub-micromolari, rendono il derivato **30r** il dichetoacido bifunzionale più attivo riportato in letteratura.

Saggi di attività biologica contro HIV-1

I composti di questa serie sono stati testati per valutare la loro capacità di inibire la replicazione dell'HIV in cellule. I composti sono stati solubilizzati in DMSO alla concentrazione di 200 mM e diluiti col mezzo di coltura. Le cellule MT-4, C8166, H9/IIIB e CEM sono fatte crescere a 37°C in un'atmosfera al 5% di CO₂ in mezzo RPMI 1640, addizionato al 10% con siero fetale di vitello (FCS), 100 IU/ml di penicillina G e 100 g/ml di streptomicina. Le colture cellulari sono testate periodicamente per verificare l'assenza di contaminazione da micoplasma con Myco Tect Kit (Gibco). I virus dell'immunodeficienza umana di tipo-1 (HIV-1, catene III B), sono ottenuti dal supernatante di cellule H9/IIIB persistentemente infettate. Le soluzioni di HIV-1 usate, hanno un titolo di 4.5×10^6 volte la dose in grado di infettare il 50% di una cultura cellulare (CCID₅₀)/ml. La titolazione dell'HIV-1 è stata effettuata in cellule C8166 attraverso il metodo di diluizione standard (diluizione 1:2), in una piastra a 96 pozzetti. Il titolo del virus è determinato con il microscopio ottico, segnando la citotossicità dopo 4 giorni di incubazione ed esprimendo il titolo virale in CCID₅₀/ml. L'attività dei composti contro la moltiplicazione dell'HIV-1 nelle cellule infettate in fase acuta è basata sull'inibizione della citotossicità virus-

indotta, nelle cellule MT-4. Il mezzo di coltura (50 μ l), contenente 1×10^4 cellule, è aggiunto in alcune flat-bottom contenenti 50 μ l di mezzo di coltura con o senza i composti testati a varie concentrazioni. Vengono quindi aggiunti 20 μ l di una sospensione di HIV-1 contenente 100 CCID₅₀. Dopo 4 giorni di incubazione a 37°C, viene determinato il numero di cellule vitali con il metodo del bromuro di 3-(4,5-dimetiltiazol-1-il)-2,5-difenil tetrazolio (MTT).

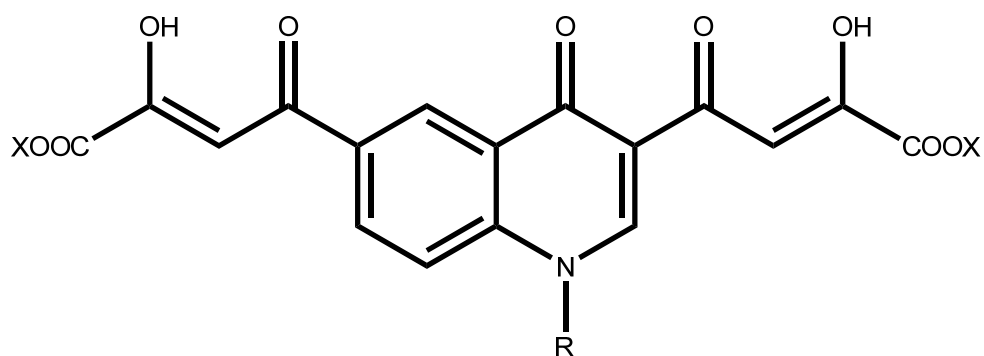
Saggi di citotossicità

La citotossicità dei composti è stata valutata su linee cellulari di *human histiocytic lymphoma* (U937) ottenute dalla “American type culture collection” (ATCC, Rockville, MD). Le cellule sono poste in una piastra a 96 pozzetti ad una concentrazione di 5×10^3 /ml nel mezzo di coltura RPMI-1640 senza rosso fenolo, addizionato con il 20% di siero fetale di vitello ed antibiotici. Dopo due ore si aggiungono i composti da testare a differenti concentrazioni (comprese tra 1 mg/ml e 200 mg/ml) e le cellule vengono incubate a 37°C per 24 ore. Quindi le cellule sono incubate a 37°C per 3 ore con 1 mg/ml di MTT (Sigma-Aldrich,

Milano, Italia). Dopo incubazione il formazano insolubile in acqua viene sciolto in isopropanolo assoluto contenente HCl 0.1 N. L'assorbanza viene misurata attraverso un lettore ELISA operante alla lunghezza d'onda di 570 nm. La citotossicità dei composti viene calcolata come la percentuale di riduzione della vitalità cellulare, per confronto con la vitalità delle cellule della coltura di controllo in assenza di farmaco. La concentrazione di farmaco richiesta per ridurre la vitalità delle cellule del 50% è detta CC_{50} .

Tabella 4 Attività enzimatica e cellulare dei derivati 30 a-d

ATTIVITÀ ANTINTEGRASICA DI DERIVATI DICHETOACIDI



anti-IN activity: IC₅₀

			ST	3'P
Comp	R	X	Mg ²⁺	Mg ²⁺
30a	H	C ₂ H ₅	77 ± 13	28
30b	H	Bz	9.1 ± 1.2	22
30c	Bz	C ₂ H ₅	0.38 ± 0.01	21
30d	Bz	H	0.016 ± 0.004	0.44

Studi di molecular modeling

Per razionalizzare le relazioni struttura–attività e fornire un plausibile meccanismo di inibizione enzimatica, il composto più attivo della serie degli acidi oxochinolinil dioxobutanoici **30r** è stato sottoposto a studi di docking usando la struttura a raggi-X 1BIS dell'integrasi (subunità B)⁷⁶. Uno ione Mg^{2+} è stato collocato nel sito attivo tra i due atomi di ossigeno dei residui Asp64 e Asp116, considerando la geometria dello ione Mg^{2+} presente nella struttura 1QS4 (subunità A) della Protein Data Bank (PDB).⁷⁷ Tenendo conto della flessibilità della proteina, il derivato **30r** è stato analizzato con una simulazione molecolare dinamica (MD) per 1 ns, in accordo con il metodo Relaxed-Complex.⁷⁸ Per gli studi di docking è stato scelto il programma AutoDock 3.0.5⁷⁹, il cui algoritmo segue la completa flessibilità del ligando ed ha dimostrato di essere in grado di rappresentare in modo soddisfacente numerose strutture cristalline complesse.⁸⁰

La proteina analizzata con la MD sembra essere molto stabile, ad eccezione dei tre loop 139-147, 165-172 e 185-196. Il primo, detto anche “loop catalitico”, è molto vicino al sito catalitico, ed è importante, data la sua

flessibilità, per l'attività catalitica dell'integrasi 81. In accordo con i precedenti studi MD il loop catalitico si muove indietro verso la regione del sito attivo⁸¹.

L'atomo di Mg^{2+} coordinato dai residui Asp64 e Asp116 resta inalterato durante tale movimento. Il docking di **30r**, realizzato attraverso un'analisi di circa 100 ps, mostra due possibili modelli di legame. In **figura 21a** è mostrato uno dei possibili binding mode del derivato analizzato: il gruppo carbossilico di una delle due catene dichetoacide del ligando chela lo ione Mg^{2+} , mentre l'altro si inserisce tra i residui Lys156 e Lys159, formando dei legami idrogeno con entrambe le catene laterali e con il carbonile dell'Asn155. Il gruppo 4-fluorobenzilico è invece posizionato in una tasca idrofobica formata dai residui del loop catalitico Tyr143, Pro142, Ile141, Gly140 e dai residui Ile60, Gln62, Val77, Val79, His114, Gly149, Val150, Ile151, Glu152, Ser153, Met154. In particolare si ha un'interazione elettrostatica favorevole tra l'atomo di fluoro del benzile con il gruppo ammidico del residuo Gln62. L'anello chinolinico della molecola instaura un'interazione ammide-anello aromatico con la catena laterale di Asn155, mentre l'ossigeno carbonilico della chinolina non è in grado di formare un legame idrogeno con la catena laterale di Thr66. Nel nostro modello Thr66, Ser153 e Met154, le cui mutazioni sembrano essere responsabili della resistenza ai DKAs, sono tutti in prossimità del ligando⁸².

Un'alternativa molto significativa a tale modello di legame è fornita dal modello in cui il metallo è chelato in modo bidentato dall'ossigeno carbonilico della chinolina e dall'ossigeno della funzione dichetoacida, mentre i gruppi ossidrilico e carbossilato della stessa catena laterale legano rispettivamente l'Asp116 e l'Asn117 (**figura 21b**). Sul lato opposto della molecola, l'altra catena dichetoacida si estende tra i residui di Lys156 e Lys159, entrando in contatto esclusivamente con l'ultimo amminoacido. Il gruppo 4-fluorobenzilico è completamente inserito nella tasca idrofobica in cui è stabilizzato da interazioni idrofobiche con le catene laterali di Ile151 e Pro142. Sebbene i due modelli di legame differiscano in alcuni dettagli, in entrambe i casi i ligandi interagiscono con l'enzima attraverso gli stessi punti di attacco:

- 1) ione metallico;
- 2) lisina 156 e/o lisina 159;
- 3) tasca idrofobica.

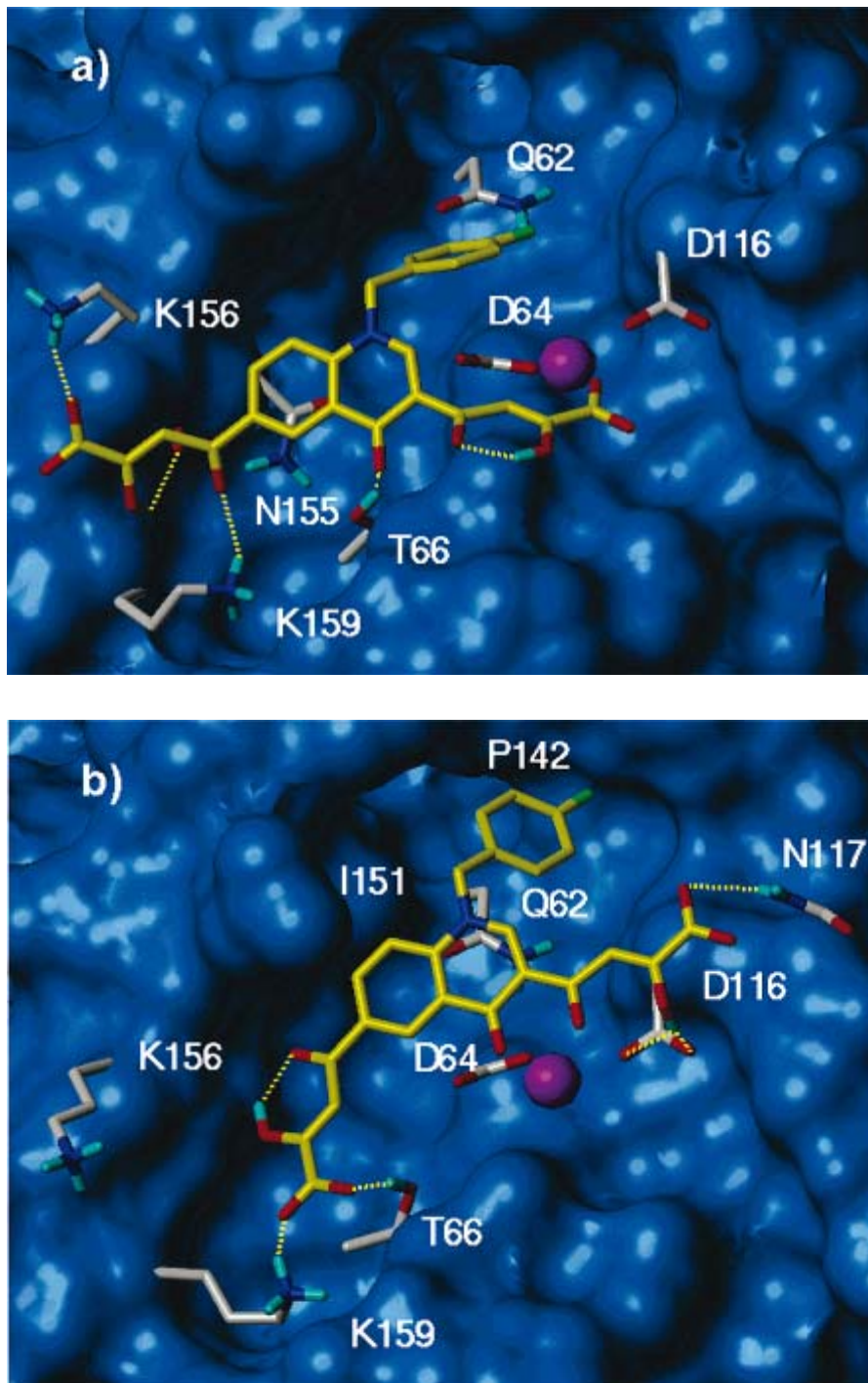


Figura 21. I due principali binding modes di 30r nel sito catalitico dell'integrasi.

La ridotta attività dei dichetoesteri rispetto ai corrispondenti di chetoacidi è dovuta alla capacità delle funzioni carbossiliche di chelare uno ione metallico e di interagire con i residui di lisina precedentemente nominati, a differenza dei dichetoesteri che legano solo blandamente questa porzione dell'enzima. Quindi per una forte interazione con l'enzima integrasi, ed in particolare con lo ione metallico, sono indispensabili le funzioni carbossiliche libere. D'altro canto, la perdita di attività osservata per i derivati non sostituiti sull'atomo di azoto dell'anello oxochinolinico rispetto a quelli N-alchilati, sia nel 3'-processing che nello strand transfer, può essere chiaramente attribuita all'assenza della catena benzilica, che nel nostro modello si inserisce nella tasca idrofobica, creando un legame più stabile. Come già riportato in precedenza⁸³ il derivato **30r** è efficace su entrambe gli step catalitici dell'integrasi, il 3'-processing e lo strand transfer, differenziandosi perciò dai dichetoacidi monofunzionali, che generalmente sono selettivi nei confronti dello strand transfer. Tenedo conto di ciò, abbiamo ipotizzato che i di chetoacidi bifunzionali (BDKAs) possano legare sia il sito accettore che il sito donatore dell'integrasi di HIV-1, essendo inibitori sia del 3'-processing che dello strand transfer, a differenza dei di chetoacidi monofunzionali che legano solo il sito accettore, essendo efficaci solo contro lo strand transfer⁸⁴. A questo punto per comprendere a fondo il meccanismo d'azione di tali derivati, abbiamo riesaminato i risultati del docking, con l'aiuto

di modelli teorici del tetramero dell'integrasi dell'HIV-1 complessata con il DNA donatore ed il DNA accettore, come riportato da McCammon e collaboratori⁸¹.

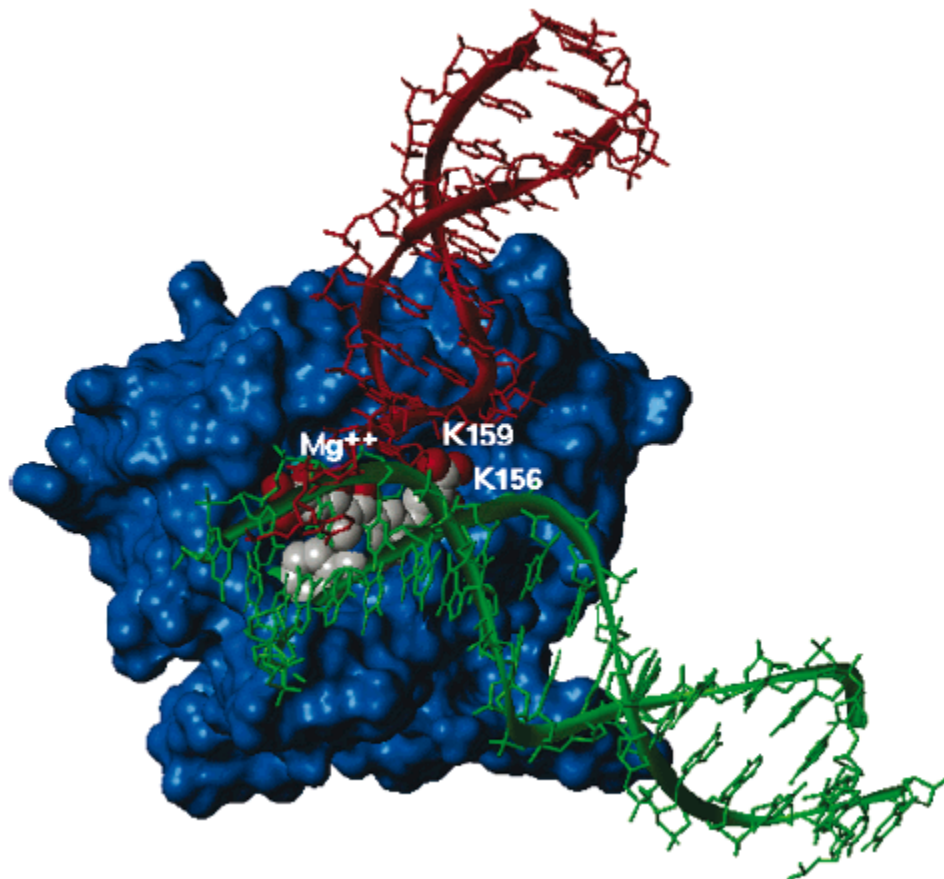


Figura 22 Modello del complesso CCD-DNA con 30r nel sito catalitico dell'integrasi.

L'immagine del composto **30r** minimizzato nel complesso CCDDNA porta alle seguenti considerazioni:

- 1) la catena dichetoacida che coordina lo ione Mg^{2+} è chiaramente importante nell'ancoraggio del ligando al sito catalitico. È noto che mentre questa porzione della molecola non sembra interferire con il legame al DNA donatore può impedire fisicamente il legame al DNA accettore. Nel nostro modello sia il DNA accettore che quello donatore hanno un contatto con il residuo Asn117 (**figura 22**). Questo amminoacido gioca un ruolo cruciale nella funzionalità enzimatica ed una sua mutazione porta ad un'integrasi difettosa soprattutto nello strand transfer⁸⁵;
- 2) l'altra funzione dichetoacida, inserendosi tra i residui di lisina 156 e lisina 159, potrebbe interferire con l'avvicinamento ed il legame del DNA donatore. Questo risultato è in accordo con i dati di photocrosslinking, che rivelano che le lisine sopra citate sono importanti per il legame al DNA donatore⁸⁶. Questo potrebbe spiegare la capacità del derivato **30d** di inibire il 3'-processing;

3) il gruppo 4-fluorobenzilico, che si inserisce nella tasca idrofobica, potrebbe ridurre la mobilità del loop catalitico ed impedire fisicamente il legame del DNA accettore. Nel modello proposto entrambe le molecole di DNA sono poste in prossimità del loop; in particolare, il DNA accettore è in stretto contatto con i residui Ile151 e Pro142, che contribuiscono al legame dell'inibitore (**figura 22**). Questi risultati confermano, in accordo con le osservazioni sperimentali, che il gruppo benzilico nei DKAs è determinante nell'inibizione dello strand transfer.

Metodi computazionali

Dinamica Molecolare

La struttura del dominio catalitico (CCD) dell'integrasi di HIV-1 è stata ricavata dalla Protein Data Bank (PDB), ref.code 1BIS¹⁰⁸. Per i nostri studi è stata presa in considerazione la catena B della struttura a raggi X. Uno ione Mg^{2+} è stato posizionato nel sito attivo tra gli atomi di ossigeno del carbossilato dei residui Asp64 e Asp116, tenendo in considerazione la geometria dello ione Mg^{2+} presente nella subunità A della struttura 1QS4 della PDB⁷¹. Per avere un'idea dello spazio conformazionale dell'enzima integrasi, è stata realizzata una simulazione MD a 1 ns, in alcuni solventi. Tutte le molecole di acqua presenti nella struttura a raggi X dell'enzima sono state allontanate e sono stati aggiunti atomi di idrogeno ai residui, considerando i residui Arg, Lys, Glu e Asp nella loro forma carica, e tutte le istidine neutre. Per rendere il sistema neutro, sono stati aggiunti due ioni Cl^- con l'aiuto del programma VEGA (versione 1.5.0). L'enzima è stato quindi solvatato con uno strato di acqua spesso 10 Å. I calcoli sono stati effettuati con il modulo DISCOVER del programma INSIGHT II,

usando il campo di forza CVFF. È stata usata una procedura multi-step. L'energia del complesso è minimizzata con 5000 step di minimizzazione a discesa rapida, seguiti da 3000 step di minimizzazione a gradiente coniugato per riordinare le molecole di acqua e i contro ioni ed eliminare ogni residua costrizione geometrica, mantenendo gli atomi di metallo dell'enzima fisso. Come struttura iniziale per lo stadio di equilibrizzazione (100ps) è stato usato il sistema minimizzato solvatato, tale stadio è seguito da un "production run" di 1 ns. Nello stadio di equilibrizzazione l'energia di minimizzazione delle catene laterali delle proteine sono state calcolate impiegando 5000 step di discesa rapida e 3000 step di gradiente coniugato. Il sistema è stato quindi scaldato gradualmente da 10 a 300 K in step da 5 ns ed è stato equilibrato a 300 K. Durante l'equilibrizzazione ed il production run, la struttura secondaria dell'enzima è mantenuta fissa. Per le interazioni di non-legame è stata considerata un'interruzione di 13 Å. Le coordinate e le energie sono state salvate ogni 10 ps, fornendo 100 diverse strutture. Il docking inoltre ha scattato in automatico un'istantanea ogni 100 ps.

Studi di docking

Il docking di **30r** nel CCD è stato realizzato usando il programma AutoDock (versione 3.0.5). L'algoritmo LGA, come effettuato nel programma AutoDock, applica un protocollo con un numero di valutazioni di energia pari a 1.5×10^{-6} , una velocità di mutazioni di 0.01, una velocità di crossover di 0.80 ed un valore di elitismo pari a 1. Per lo studio di **30r** sono stati realizzati 50 docking indipendenti in differenti istantanee MD dell'enzima. I risultati che differiscono per meno di 1.5 Å per ogni posizione (metodo dei minimi quadrati), sono sempre raggruppati e rappresentati dal risultato con energia libera di legame più favorevole. I complessi ottenuti sono minimizzati energeticamente usando 3000 step dell'algoritmo a discesa rapida, permettendo solamente al ligando e agli atomi della catena laterale della proteina di rilassarsi entro un raggio di 5 Å intorno al ligando. L'ottimizzazione geometrica è stata realizzata impiegando il programma DISCOVER con il campo di forza CVFF.

Ligand setup

La struttura di **30r** è stata generata dai frammenti standard della libreria del software SYBYL versione 7.0. L'ottimizzazione geometrica è stata realizzata con il minimizzatore SYBYL/MAXIMIN2 applicando l'algoritmo BFGS con una convergenza dei criteri di 0.001 kcal/mole ed impiegando il campo di forza TRIPOS. Le cariche atomiche parziali sono state assegnate usando i formalismi di Gasteiger e Marsili, come impostato nel pacchetto SYBYL. Il derivato **30r** è stato modellato nella sua forma cheto-enolica con la carica negativa sul gruppo carbossilato. Sono state specificate sei possibili torsioni: due intorno all'anello chinolinico, due a livello del gruppo benzilico e due che consentono la rotazione dei gruppi ossidrilici.

Protein setup

La struttura della proteina è stata settata per il docking come segue: sono stati rimossi tutti gli atomi di idrogeno non polari e sono state assegnate le cariche atomiche parziali secondo Kollman. Sono stati aggiunti dei parametri di solvatazione al complesso proteina-DNA usando l'utilità ADDSOL del

programma AutoDock. Le mappe delle griglie sono state calcolate con l'ausilio di AutoGrid. Le griglie scelte sono abbastanza larghe in modo da includere una parte significativa della proteina intorno al sito catalitico. In tutti i casi sono state usate griglie di 61 x 61 x 61 punti con uno spaziatura griglia-punto di 0.375 Å. Il centro della griglia è stato fatto coincidere con lo ione Mg^{2+} nel sito attivo del dominio catalitico.

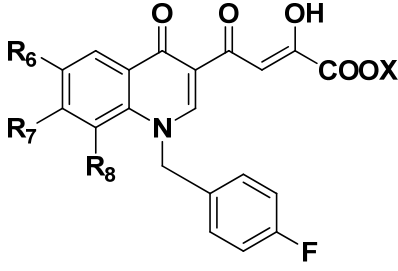
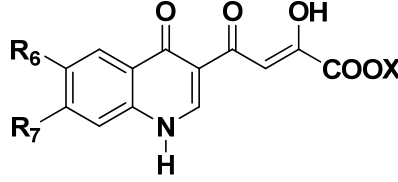
Parte IV

Progettazione e sintesi di inibitori dichetoacidi monofunzionali (DKA)

Visti i promettenti risultati relativi ai derivati chinolin di chetoacidi bifunzionali, successivamente è stata progettata una nuova serie di derivati chinolonici monofunzionali, questa volta, selettivi verso lo strand trasfer, allo scopo di fornire nuove conoscenze circa il meccanismo catalitico di IN. Pertanto sono state studiate le relazioni struttura attività relativamente alla serie di derivati chinolin dichetoacidi **31a-i**, **32a-i**, **33a,b** e **34a,b** in cui è stata sostituita la catena dichetoacida di 4, responsabile del legame al DNA donatore con sostituenti più piccoli, nelle posizioni 6, 7, e 8 dell'anello chinolinico. Questi composti dovrebbero esibire un ridotto legame al DNA donor sul quale si attacca il DNA provirale per la presenza di piccoli sostituenti che limitano la capacità di

formare legami ad H. Pertanto le modifiche strutturali apportate sulla serie forniscono nuovi derivati dichetoacidi con incrementata selettività verso lo ST.

Tabella 5 Attività enzimatica e cellulare dei derivati 31 a-i, 32 a-i, 33 a-b, 34 a-b

Anti-IN activity: IC ₅₀												
<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  <p>31a-i, 32a-i</p> </div> <div style="text-align: center;">  <p>33a-b, 32a-b</p> </div> </div>												
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>R₆, R₇, R₈ = H, F, Cl, 1-pyrrolidinyl X = H, C₂H₅</p> </div> <div style="width: 45%;"> <p>R₆, R₇ = Cl X = H, C₂H₅</p> </div> </div>												
Compd.	R ₆	R ₇	R ₈	X	ST		3'-P	antiviral activity				
					Mg ^{2+b}	Mg ^{2+c}	Mg ^{2+c}	CC ₅₀ ^d	EC ₅₀ ^e	SI ^f	T-KI ^g	
31a	H	H	H	C ₂ H ₅	0.67	0.79, 0.44	135	99	34.8	2.8		
31b	F	H	H	C ₂ H ₅	0.58	0.23, 0.44	>333	>200	>50	-		
31c	H	F	H	C ₂ H ₅	1.3	1.2, 2.9	>333	>200	>50	-		
31d	H	H	F	C ₂ H ₅	26	5.2	>333	>200	>50	-		
31e	Cl	H	H	C ₂ H ₅	21	3.0	>333	nt ^g	nt			
31f	H	Cl	H	C ₂ H ₅	14	3.6	>333	nt	nt			
31g	H	H	Cl	C ₂ H ₅	32	5.2	>333	>200	>50	-		
31h	Cl	Cl	H	C ₂ H ₅	3.9	1.8	135	72	>50	-		
31i	H	Py ^b	H	C ₂ H ₅	2.3	2.7	110	>200	4.1	48.8		
32a	H	H	H	H	0.040	0.06, 0.16	14	>200	1.17	>171	0.052	
32b	F	H	H	H	0.030	0.14	2.3	>200	>50	-	0.065	
32c	H	F	H	H	0.020	0.050	4.4	>200	>50	-	0.071	
32d	H	H	F	H	0.018	0.16	16	>200	>50	-	0.046	
32e	Cl	H	H	H	0.028	0.40	6.4	>200	46.1	4.3		
32f	H	Cl	H	H	0.019	0.11	4.0	>200	>50	-	0.015	
32g	H	H	Cl	H	0.023	0.51	44	>200	>50	-		
32h	Cl	Cl	H	H	0.033	0.075	11, 34	156	3.2	62.5		
32i	H	Py ^b	H	H	0.028	0.18	14, 9.0	>200	0.17	>1176	0.031	
33a	Cl	H		C ₂ H ₅	3.3	2.1	120	nt	nt			

33b	H	Cl	C ₂ H ₅	>111	224	>333	nt	nt
34a	Cl	H	H	>111	56	166	nt	nt
34b	H	Cl	H	37	17	>333	nt	Nt

^aInhibitory concentration 50% (μM) determinata dalla curva dose-risposta. ^bEsperimenti fatti usando il saggio dell' high throughput electrochemiluminescent (HTECL) (saggio ST in presenza di Mg^{2+}). ^cEsperimenti condotti su gelin presenza di Mg^{2+} . ^dCytotoxic concentration 50% (μM). ^eEffective concentration 50% (μM). ^fSelectivity index = $\text{CC}_{50}/\text{EC}_{50}$. ^gnt: not tested. ^hPy = 1-pyrrolidinyl. ⁱQuesto dato è riferito a EC_{95} . ⁴T- K_i è la K_i teorica (μM) predetta dal software AutoDock.

Saggi di attività biologica verso HIV-1

I composti neo sintetizzati sono stati testati per la loro capacità di inibire lo strand trasfer attraverso vari saggi in vitro. I derivati sopra riportati sono stati testati in vitro per lo ST in presenza di magnesio utilizzando il saggio di elettrochemiluminescenza ad alta portata (high throughput electrochemiluminescenz *HTECL*) per ottenere i valori di IC_{50} . I composti poi sono stati anche testati per lo ST ed il 3'-P usando saggi su gel in presenza di Mg^{2+} . I DKA neo sintetizzati hanno mostrato un'alta selettività nei confronti dello ST confermando così l'ipotesi che rimuovendo la porzione dichetoacidica in posizione 6 dell'anello chinolonico si potessero ottenere inibitori selettivi per lo ST. I derivati 32a-i sono più potenti anche dei corrispondenti esteri 31a e gli 1-p-F-benzyl chinolinoni sostituiti (**31e,f** e **32e,f**) sono sensibilmente più attivi della loro controparte non sostituita (**33 a,b** e **34a,b**). Questi trend combaciano

con i risultati ottenuti per la serie dei BDKA. In particolare il ruolo rilevante giocato dalla porzione p-F-benzilica in posizione 1 dell'anello chinolinonico è stato confermato dal binding ottimale sul target biologico. I composti **32a-1** si sono rivelati essere i più potenti inibitori mostrando dei valori di IC₅₀ che variano tra i 18 ed i 40 nM per lo step dello ST (**tabella 5**). Il composto più attivo della serie **32d** ha mostrato un valore di IC₅₀ verso lo ST di 18nM. L'introduzione di un alogeno (F o Cl) nelle posizioni 6, 7 oppure in 8 ai composti **32b-g** ha messo in evidenza dei valori di attività antivirale comparabili rispetto al composto **32a** non sostituito. Tali valori sono influenzati dalla posizione in cui si trova l'alogeno. Infatti le attività decrementano nel seguente ordine: 8-F(Cl) ≈ 7-F(Cl) > 6-F(Cl) e 7-Cl > 8-Cl > 6-Cl. L'introduzione di un secondo atomo di cloro nella struttura del potente inibitore **32f** ha prodotto **32h** che si è mostrato 1.5 volte meno potente di **32f**. Similarmente la sostituzione di un atomo di cloro di **32f** con un anello pirrolidinico ha prodotto **32i** che si è mostrato 1.5 volte meno potente di **32f**. In conclusione le attività dei composti con un sostituito in posizione 7 decrementano nel seguente ordine: Cl > F > 1-pyrrolidinyl > H (**tabella 5**). I composti **31a-i**, **32a-i**, inibiscono l'integrasi con potenza simile in presenza di Mg²⁺ o di Mn²⁺.

Saggi di citotossicità

La citotossicità e le attività antivirali dei composti **31a-i**, **32a-i**, **33a,b**, **34a,b** sono rappresentate in **tabella 5**. In generale tutti i composti testati sono caratterizzati da una bassa citotossicità quando testati verso la linea cellulare human histiocytic lymphoma (U937) e mostrano un CC_{50} che varia tra 72 e 200 μM . Tra i composti testati, i derivati **31a,i**, e **32a,e,h,i** mostrano un $EC_{50} < 50$ μM quando analizzati sulle cellule H9/HTLVIIIB infettate. In particolare, **31i** e **32a,h,i**, sono attivi nel range del micromolare con S.I. che varia tra 48.8 e 1176. Il più potente derivato di questa serie è risultato essere **32i** ($EC_{50} = 0.17$ μM e S.I. > 1176).

Molecular modeling

Pertanto per esplorare i possibili binding mode dei derivati-neosintetizzati, nel sito attivo di IN, considerando che l'inibitore 32a selettivo per lo strand trafer si lega al core domain di IN senza interferire con il DNA binding virale, è stato costruito un modello ternario di integrasi contenente Mg^{2+} e DNA virale. A tale scopo è stata adoperata l'ultima versione di Autodock (vs 4.0) con una funzione di scoring implementata per il calcolo dell'energia libera di legame.

Generazione del complesso ternario con due metalli e DNA virale

La struttura cristallina di HIV-IN CCD risolta da Maignan et al⁸⁷. (PDB accession: 1BL3_C) è stata usata per la costruzione di tale modello. In essa si evince un loop catalitico flessibile completamente risolto ma in confronto a 1BIS_B⁷⁶, che pure presenta un loop catalitico completamente risolto permette la coordinazione con un secondo metallo attraverso la rotazione della catena laterale di E152. Come mostrato in **figura 4a**, la sovrapposizione strutturale dei $C\alpha$ di 1BL3_C ed 1BIS_B mostra che oltre alle ben note differenze strutturali relative alla conformazione del loop flessibile, le due strutture divergono per l'orientazione di una elica $\alpha 4$ che appartiene al residuo E152. In 1BIS_B,

l'amminoacido E152 è più lontano da D64 e D116 rispetto alla struttura di 1BL3_C, ed una semplice rotazione della catena laterale di E152 potrebbe non seguire la coordinazione di un secondo metallo (in 1BIS_B, le distanze tra i C α di E152 e D64 C α e di E152 e D116 sono di 13.3 e 9.7 Å, rispettivamente). Pertanto, 1BL3_C è sicuramente più appropriata di 1BIS_B per la costruzione di un modello di IN con due metalli.

Sovrapponendo le due polinucleotidil polimerasi ASV IN (PDB accession: 1VSH)⁸⁸ e Tn5 transposase (PDB accession: 1MUS) sui C α della triade catalitica risulta che la posizione dei due metalli e la loro coordinazione attraverso la triade catalitica è molto vicina. La struttura di 1BL3_C contiene già un metallo nella cavità catalitica ed il posizionamento del secondo è stato effettuato secondo la procedura descritta nell'Experimental Section. Come risultato di ciò i due metalli sono distanti tra loro 3.8 Å: il primo è coordinato da D64 e D116 mentre il secondo da D64 e E152.

Come sopra menzionato, nessuna struttura di IN risolta ai raggi X mostra il DNA legato all'enzima, sebbene siano stati proposti differenti modelli teorici^{89-91,82}. Recentemente un'importante fonte di informazioni sulla corretta collocazione del DNA nel dominio catalitico di IN è stata identificata nella struttura cristallina del complesso di Tn5 transposase/DNA⁹² Tn5 transposasi è un enzima appartenente alla famiglia delle polinucleotidil transferasi cui

appartiene anche IN e tra esse sono evidenti molte similarità⁹³. Come IN, Tn5-transposasi contiene un motivo DDE altamente conservato e richiede uno ione Mg^{2+} o Mn^{2+} per la sua attività catalitica. Un tentativo sperimentale di dimostrare l'alta similarità tra i due enzimi è stato fatto attraverso l'identificazione di inibitori di IN da uno screening in vitro usando Tn5 transposase come target⁹⁴⁻⁹⁵.

La posizione del DNA virale nel sito attivo di IN, è stato studiato da Wielens et al.⁹⁶ che ha costruito un modello teorico del complesso di integrazione (PDB accession: 1ZA9), usando la struttura cristallina di Tn5 transposase come template. In questo modello IN è complessato con il DNA virale 3'-processato. Pertanto le coordinate di tale DNA virale sono state adoperate per la costruzione del nostro modello. Per refinire tale modello ternario e considerare la flessibilità della proteina per i successivi esperimenti di docking è stata effettuata una dinamica molecolare per 3 ns in solvente esplicito. Le proprietà termodinamiche sono state monitorate durante la simulazione e fino a convergere a valori stabili. L'analisi della struttura secondaria inoltre ha rivelato la presenza di motivi secondari altamente conservati ed invece la mobilità di alcune porzioni della proteina soprattutto confinate nella regione dei loops.

Studi di docking

Per comprendere il binding mode del più potente inibitore della serie 32i in IN catalytic domain, sono stati condotti esperimenti di docking in un ensemble di IN/DNA/metal snapshots raccolti ogni 300 ps durante una simulazione di 3 ns in accordo al relaxed-complex method⁷⁸.

I docking di **32i** negli MD snapshots hanno rivelato tre conformazioni relativamente molto vicine. Tra queste, la prima, in cui in tutti gli snapshots adoperati presentano la più bassa energia di binding ($\Delta G_{bind} = -10.2$ kcal/mol), è stata caratterizzata anche da un maggior numero di interazioni inibitore-enzima rispetto alle altre due conformazioni di binding. In questa posa (**figura 5a**), l'inibitore si dispone parallelamente all'asse del DNA occupando la regione della proteina costituita dai residui C65, T66, H67, L68, E92, N155 e K159. Tra questi, singole mutazioni di C65 e T66 sono state sufficienti a conferire resistenza verso gli inibitori di IN⁹⁷⁻⁹⁸.

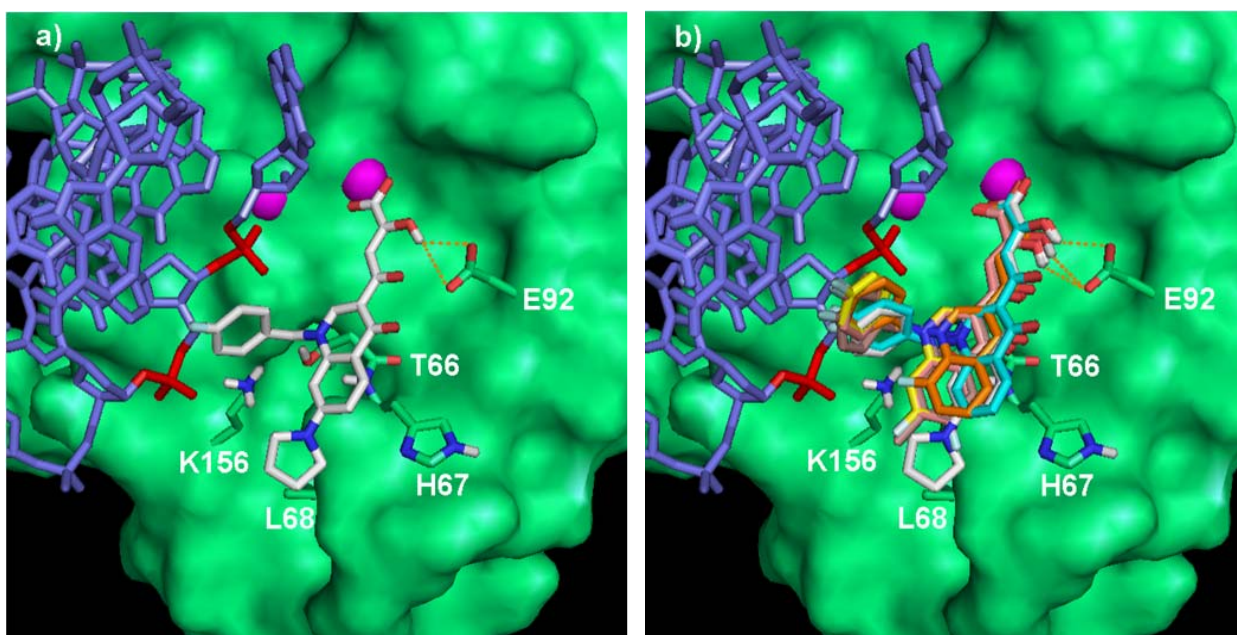


Figura 23. a) Composto **32i** (in bianco) è legato nel sito attivo di HIV-1 IN rappresentata come una superficie di Connolly verde. Il DNA virale è blu con due gruppi fosfato, in rosso che sono disposti vicino al ligando. I residui di IN importanti per il binding dell'inibitore sono mostrati in sticks verdi. Gli idrogeni polari sono stati rimossi per chiarezza. I legami ad H sono mostrati come dashed lines arancioni. I due metalli di Mg^{+2} sono rappresentati come sfere in viola b) Sovrapposizione di **32a** (ciano), **32b** (rosa), **32c** (giallo), **32d** (arancio) e **32f** (marrone) nei loro migliori binding modes.

L'inibitore partecipa attraverso i due ossigeni del carbossilato alla coordinazione dello ione metallico in concomitanza con i residui D64 e D116 del sito catalitico, in accordo anche ai dati SAR secondo i quali i DKA presentano valori di attività inibitoria più alti rispetto ai corrispondenti esteri. La coordinazione del secondo metallo è evidentemente impedita dalla presenza del DNA virale. Il gruppo ossidrilico enolico di **32i** forma inoltre legami ad H con il residuo E92. L'anello chinolonico giace sotto il legame ammidico tra H67-T66 e

forma un interazione π - π stacking off-centered con l'anello imidazolico di H67 (la distanza tra i centroidi degli anelli è 4.9 Å). L'anello elettrone-deficiente *p*-F-benzilico punta verso il DNA virale formando favorevoli interazioni di stacking. Questo potrebbe spiegare la caduta dell'attività quando l'anello *p*-F - benzilico è assente (vedi composti **34a** e **34b**). L'MD snapshot del complesso IN/DNA/metallo che ha mostrato la miglior binding pose per **31i** è stato scelto per tutti i docking degli inibitori attivi **32a-d** e **32f**. Le pose di binding con i più bassi ΔG_{bind} si sono mostrate sovrapponibili a quella descritta per **31i**. Inoltre anche le principali interazioni con l'enzima ed il DNA virale si sono conservate. Tuttavia ad esclusione di **32a**, che ha mostrato un binding mode identico a quello di **32i**, un leggero spostamento dell'anello chinolonico è stato osservato per i composti **32b-d** e **32f**. Ciò è stato ascritto alla presenza degli alogeni sull'anello chinolonico. Infatti quando l'atomo di fluoro è posizionato in 6, composto **32b**, il piccolo incremento di attività sembra essere dovuto ad un incrementato π - π stacking off-centered con la catena laterale di H67. Al contrario, quando il fluoro è in posizione 7 e 8 (composti **32c** e **32d**), si forma un legame ad idrogeno con la catena laterale di K159. Il composto **32f** che presenta un atomo di cloro in posizione 7 dell'anello chinolonico, presenta lo stesso comportamento di **32c**. Questi risultati sono in accordo con le SAR che

mostrano un incremento di attività inibitoria verso IN di **32c**, **32d** e **32f** comparati a **32a**.

I valori della K_i teorica (T- K_i) predetti da Autodock per ognuno dei composti dockati, può essere comparata ragionevolmente ai dati di inibizione sperimentali, che suggeriscono affinità nel range nanomolare. I valori di T- K_i ottenuti per i composti dockati sono compresi tutti nel range di 15- 71 nM, concordando con i valori di IC_{50} sperimentali (HTECL assay), che variano tra i 18 to 40 nM. I binding mode proposti per **32a-d**, **32f** e **32i** sono in accordo con gli esperimenti di base di Schiff e cross-linking che hanno indicato **32a** come specifico inibitore dello ST che si dispone in IN CCD con un'orientazione tale da non interferire con il binding del DNA virale ed il 3'-P⁹⁹. Questi esperimenti anche suggeriscono che il ligando non si può legare in prossimità di Q148, che appartiene al loop flessibile né in prossimità di K156 e K159, conosciuti per interagire con il DNA virale, ma essi invece sono capaci di bloccare il sito del DNA accettore, spiegando così l'inibizione selettiva dello ST¹⁰⁰. I nostri studi di docking sono in perfetto accordo con questa ipotesi come illustrato in **figura 23**. Infatti inserendo la predetta conformazione di **32i** nel modello a due metalli complessato con DNA virale ed umano come descritto nella sezione dei metodi computazionale per la costruzione del modello quaternario, risulta che la branch dichetoacidica che occupa il centro del sito

catalitico potrebbe interferire solo con il DNA umano. Questo è il motivo per cui si osserva la selettiva inibizione dello ST process, con il coinvolgimento di un solo metallo allo stesso modo importante nel processo di inibizione dello strand transfer.

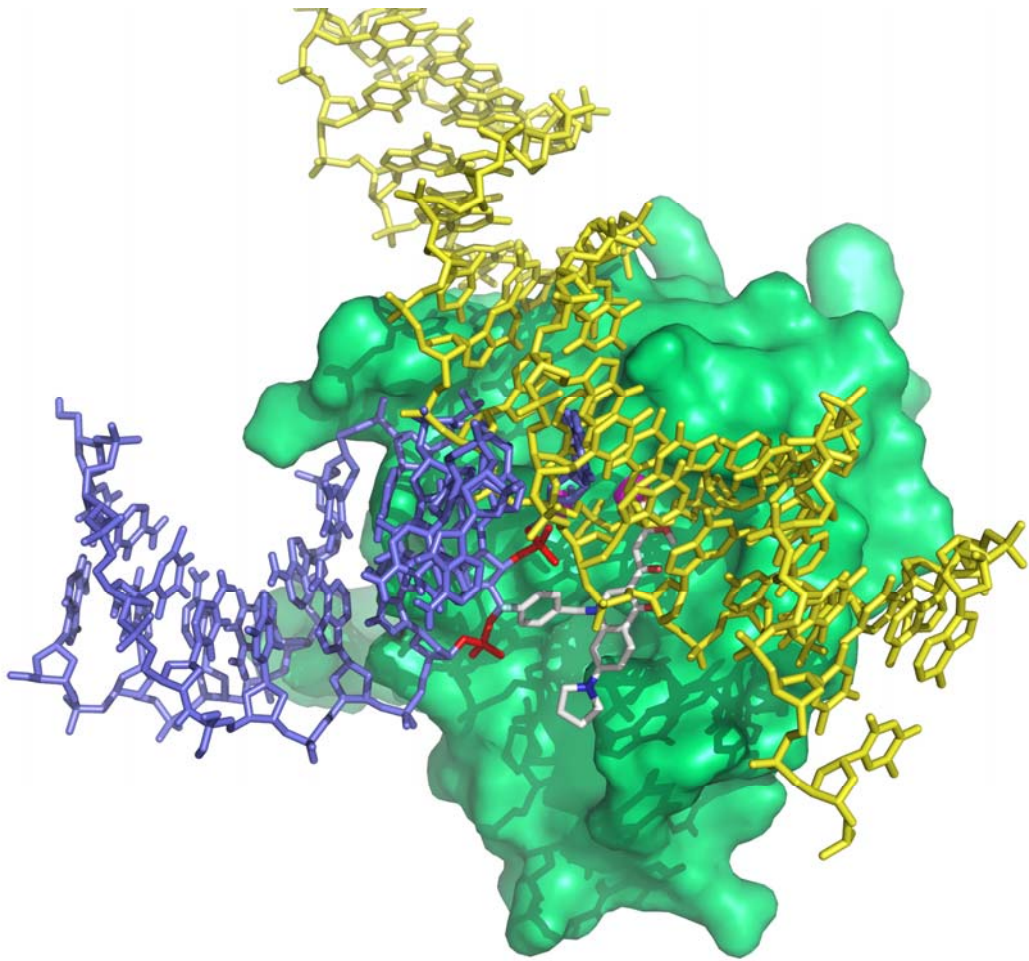


Figura 24. Modello a due metallic di HIV-1 IN CCD complessato con entrambi DNA virale (blu) e DNA umano (giallo). L'inibitore 32i è visibile ed è rappresentato da sticks bianchi.

Pertanto, partendo da questo modello, di IN, complessato con DNA virale ed umano è stato possibile spiegare il differente meccanismo d'azione dei DKA e dei precedentemente considerati BDKA⁷⁵. Precedenti studi di docking hanno mostrato che per i BDKA , una catena laterale DKA coordina lo ione metallico Mg^{2+} mentre l'altro motivo DKA si inserisce tra le catene laterali dei residui di K156 e K159⁷⁵. Esperimenti di cross-linking hanno confermato l'interazione dei BDK con le sopra menzionate lisine⁹⁹. Differentemente dai monofunzionali DKAs, l'interazione dei BDK con K156 e K159 potrebbe fisicamente ostacolare il legame con il DNA virale, provocando una inibizione non selettiva di entrambe le reazioni dello 3'-P e ST.

Metodi computazionali

Generazione del modello IN CCD complessato con DNA virale ed umano

La struttura ai raggi X di HIV-1 risolta da Maignan et al. (PDB ID 1BL3_C)⁸⁷ è stata usata come base per il modelling di IN CCD complessato con due metalli per la generazione di un modello adatto, attraverso la sovrapposizione di 1BL3_C sulla struttura cristallina di Tn5 transposase (PDB accession: 1MUS) usando un rigid-body least-squares fit basato sulle coordinate dei C α della triade catalitica (D64, D116, E152 per 1BL3_C e D97, D188, E326 for 1MUS). La catena laterale di E152 in 1BL3_C, è stata ruotata manualmente per sovrapporla a quella di E326 nella struttura cristallina di 1MUS .

Un secondo ione magnesio è stato collocato nella stessa posizione dello ione coordinato da D97 ed E326 in 1MUS.

Il modello di IN-DNA recentemente sviluppato da Wielens et al.⁹⁶ (PDB accession: 1ZA9) è stato usato per posizionare il DNA virale nel nostro modello di IN con due metalli. Tale complesso di HIV-1 IN CCD (catena A di 1BL3) con DNA virale ed umano (catene X ed S, rispettivamente) è stato isolato. Il

modello costruito da noi con due metalli è stato sovrapposto su quello di IN-DNA di Wielens usando Align Structures routine di BIOPOLYMER module del software SYBYL 7.3¹⁰¹. Il DNA virale è stato introdotto nel CCD IN con due metalli per produrre un complesso ternario. Successivamente è stato poi collocato nel modello anche il DNA umano per produrre un modello quaternario da utilizzare per un'analisi *a posteriori* dei possibili meccanismi d'azione degli inibitori DKA e BDKA .

Simulazioni di dinamica molecolare

Per rimuovere alcuni clashes sterici tra il DNA virale e IN CCD e per conferire flessibilità al sistema per gli esperimenti successive, una simulazione dinamica a 3 ns MD del complesso ternario è stata effettuata usando SANDER module di AMBER9¹⁰². Una rappresentazione delle coordinate atomiche di tutto il sistema è stata ottenuta, utilizzando il force field di Cornell et al.¹⁰³ che ha assegnato i parametri alla proteina e alle parti nucleiche del sistema. I residui di Arg, Lys, Glu e Asp sono stati mantenuti carichi mentre i residui di His sono stati considerati neutri.

Il complesso di IN-DNA complex è stato inserito in una box di molecole di acqua (7349 molecole) di TIP3P¹⁰⁴ e neutralizzato con l'aggiunta di 15 ioni Na⁺. Il complesso iniziale è stato dapprima minimizzato con 1500 steps di steepest descendent, seguiti da 1100 steps di conjugate gradient allo scopo di ottimizzare l'orientazione delle molecole di acqua, dei controioni e di tutti gli idrogeni. Durante la minimizzazione, la proteina, il DNA e gli ioni Mg²⁺ sono stati fissati. Successivamente una minimizzazione energetica delle catene laterali e dei loops è stata condotta attraverso 2300 steps di steepest descendent seguiti da 2200 steps di conjugate gradient mantenendo il backbone della proteina ed il DNA bloccati.

Successivamente una fase di equilibratura è partita secondo il seguente protocollo:

- 60 ps di equilibratura a volume e temperatura costanti nel quale la temperatura del sistema è stata gradualmente aumentata da 100 K a 300 K, con una restraint force di partenza di 50 kcal/mol/Å² sul backbone della struttura secondaria della proteina e 500 kcal/mol/Å² sul DNA backbone e sugli ioni Mg²⁺. Mentre la temperatura aumentava, la restraint force sul backbone della struttura secondaria della proteina è stata gradualmente decrementata a 0.1 kcal/mol/Å², mentre le constraints sul DNA backbone e sugli ioni Mg²⁺ sono state mantenute.
- 40 ps di density equilibration in NPT ensemble a 300 K con una restraint force di 5 kcal/mol/Å² sul backbone della struttura secondaria della proteina che gradualmente è stata decrementata a zero, mentre le constraints sul DNA e sugli ioni Mg²⁺ ions sono state coneservate.
- 3ns di produzione è stata lanciata a temperatura costante di 300 K ed a pressione costante di 1atm. Lo SHAKE method¹⁰⁵ è stato applicato come constraint su tutti i legami covalenti coinvolgendo

gli atomi di idrogeno. Durante la simulazione, il nonbonded cutoff è stato settato a 10 Å. Nessuna constraint nè restraint è stata applicata al complesso ternario durante la produzione per seguire la completa flessibilità del sistema. L'analisi delle traiettorie è stata condotta usando PTRAJ module di AMBER software package.

Simulazioni di docking

Docking dei ligandi in IN CCD sono stati condotti con AutoDock software package (version 4.0) implementato attraverso una graphical user interface AutoDockTools (ADT 1.4.5)¹⁰⁶.

Le strutture dei ligandi sono state dapprima generate attraverso la standard fragment library di SYBYL software version 7.3. Le ottimizzazioni geometriche sono state condotte con SYBYL/MAXIMIN2 minimizer attraverso l'applicazione dell'algoritmo BFGS (Broyden, Fletcher, Goldfarb and Shannon)¹⁰⁷ con un criterio di convergenza di 0.05 kcal/mol implementato attraverso TRIPOS force field. Successivamente, i composti costruiti ed il complesso ternario sono stati convertiti negli AutoDock format files usando ADT. Tutti i composti sono stati modellati nella forma anionica come specie prevalente, a pH fisiologici. La

docking box area è stata costruita usando come suo centro lo ione metallico coordinato con D64 ed D116, tanto larga da coprire l'intera superficie della triade catalitica. Successivamente griglie di $85 \text{ \AA} \times 85 \text{ \AA} \times 85 \text{ \AA}$ distanziate 0.375 \AA sono state calcolate intorno all'area di docking per tutti gli atom types del ligando usando AutoGrid4. Per ciascun ligando sono stati condotti 50 run di docking. Ogni docking ha consistito di 2500000 energy evaluations usando il Lamarckian genetic algorithm local search (GALS) method. The GALS method valuta una popolazione di possibili soluzioni di docking e propaga quelle migliori da una generazione all'altra. È stata poi adoperata una ricerca locale a bassa frequenza, in accordo al metodo di Solis e Wets, per ottenere una soluzione finale in una situazione di minimo locale. Tutti i docking sono stati condotti con una population size di 150 e 300 rounds di Solis e Wets local search sono state applicate con una probabilità di 0.06. Una mutation rate di 0.02 ed un crossover rate di 0.8 sono stati usati per generare nuovi docking trials per le generazioni successive. I risultati sono stati clusterizzati sulla base della root-mean square deviation ($\text{rmsd} = 1 \text{ \AA}$) tra le coordinate cartesiane degli atomi del ligando e sono state raggruppate sulla base della energia libera di legame (ΔG_{bind}). I complessi ligando-IN-DNA-metallo sono stati poi minimizzati usando 5000 steps di steepest descent con TRIPOS come campo di forza. La minimizzazione è stata

condotta mantenendo gli atomi del backbone fissi mentre le catene laterali della proteina e del ligando libere di muoversi.

Conclusioni

La sindrome da immunodeficienza acquisita (AIDS), causata dal virus HIV-1, rimane ancora un problema irrisolto. Dopo anni di intenso lavoro, sono stati scoperti ed introdotti nella pratica clinica numerosi inibitori della trascrittasi inversa e della proteasi e da pochissimi anche inibitori della fusione.

Però, a causa della resistenza sviluppata dal virus HIV-1, tutte le terapie farmacologiche intraprese per inibire l'uno e l'altro enzima, si sono rivelate sfortunatamente inefficaci nei trattamenti a lungo termine. Per questo motivo si è pensato di combinare, in una unica terapia farmacologica, sia gli inibitori della proteasi che quelli della trascrittasi inversa: visto che in questo caso si tende ad inibire il ciclo di replicazione virale bloccando due diverse tappe del processo, la terapia si è rivelata molto più efficace. Dopo la scoperta dell'enzima integrasi, numerosi studi farmaceutici sono stati rivolti alla sintesi e alla progettazione di molecole capaci di inibire il target in questione. In sinergia con le terapie farmacologiche già utilizzate, l'utilizzo di queste nuove molecole ha potenziato l'effetto inibitorio nei confronti della replicazione virale.

Il lavoro presentato in questa tesi riguarda la sintesi e la progettazione di nuovi inibitori dell'integrasi a struttura quinolinonil dicheto acida bi-funzionali

attivi su entrambi gli step dello strand transfer e del 3'P e di derivati chinolinonil dicheto acidi mono- funzionali selettivi verso lo strand transfer. I derivati sintetizzati hanno mostrato un'elevata attività anti-HIV-1, del tutto paragonabile o superiore a quella di composti molto promettenti descritti in letteratura, rivelandosi attivi nei confronti di cellule infettate. Nell'ambito della serie dei chinolin di-chetoacidi bi-funzionali **30a-d**, risultati importanti sono stati ottenuti saggiando il composto **30d** che ha mostrato un'ottima azione inibente nei confronti dell'integrasi, sia sul 3'-processing ($IC_{50} = 0.44\mu M$) che sullo strand transfer ($IC_{50} = 0.0016 \mu M$), risultando dalle 17 alle 540 volte più attivo di alcuni dichetoacidi bifunzionali riportati in letteratura. Gli esperimenti di docking sono stati indispensabili per comprendere meglio l'interazione del derivato **30d** con l'enzima e ci hanno permesso di ipotizzare un suo possibile meccanismo d'azione. Esso inoltre è attivo anche nei confronti delle cellule infettate da HIV-1 ($EC_{50} = 4.29 \mu M$), risulta scarsamente citotossico $CC_{50} > 200 \mu M$ e mostra un indice di selettività > 47 . Successivamente la serie dei chinolin di-chetoacidi mono-funzionali **31a-i**, **32a-i**, **33a,b** e **34a,b** in cui è stata sostituita la catena dichetoacida di 4, responsabile del legame al DNA donatore con sostituenti più piccoli, nelle posizioni 6, 7, e 8 dell'anello chinolinico ha mostrato selettività per lo strand transfer. Tra tutti i composti **32i** ha mostrato un'ottima azione inibente selettiva nei confronti dell'integrasi, in particolare sullo

strand transfer 0.028 μM vs 9.0 μM del 3'P. Anche in questo caso gli studi computazionali hanno permesso di comprendere come tali inibitori ed in particolare il più potente della serie l'inibitore **32a** possa essere selettivo per lo strand transfer legandosi al core domain di IN senza interferire con il DNA binding virale. Tali composti potrebbero essere considerati degli ottimi derivati sui quali investire per una potenziale applicazione terapeutica.

Bibliografia

1. Robert Gallo. *A caccia di virus*. Milano, Rizzoli, 1992.
2. Campbell, G. R. et al. (2004) The glutamine-rich region of the HIV-1 Tat protein is involved in T-cell apoptosis. *J. Biol. Chem.* **279**, 48197-48204
3. Morgan, D. et al. (2002) HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries? *AIDS* **16**, 597-6032
4. Montessori V. et al. Adverse effects of antiretroviral therapy for HIV infection *CMAJ* January 20, 2004; 170
5. Grmek, M.D., *AIDS, storia di un'epidemia attuale*, Laterza, Roma-Bari 1989.
6. Montagnier, L. e Daudel, R., *Aids*, Il Saggiatore, Milano 1996.
7. UNAIDS 2008 global report.
8. Cassuto, J.P.; Pesce, A.; Quaranta, J.F., *AIDS e infezione da HIV*, Masson, Milano 1995.
9. Letvin, M.D. et al. Induction of AIDS-like disease in macaque monkeys with T-cell tropic retrovirus STLV-III. *NL. Science*, **230**, 4721, 71-73

10. Gamble T.R., Yoo S., Vajdos F.F., Hill C.P. *Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein*. Science, 278, 849- 853, 1997.
11. Gibellini D. et al. *Selective up-regulation of functional CXCR4 expression in erythroid cells by HIV-1 Tat protein*. Clin. Exp.Imuunol. 131: 428-435, 2003.
12. Barre-Sinoussi F. et al. *Isolation of a Tlymphotropic 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. Kohlstaedt L.A., Wang J., Friedman JM., Rice PA., Steitz TA. *Crystal structure at 3.5 A resolution of HIV-1 reverse trascriptase complexed with an inhibitor*. Science, 256, 5065, 1783-90. 1992.
15. Rice P. et al. *Retroviral integrases and their cousins*. Curr Opin Struct Biol. 6(1):76-83, 1996.
16. Hernandez L.D. et al. White JM. *Virus-cell and cell-cell fusion*.Annu. Rev. Cell. Dev. Biol. 12:627-61, 1996.
17. Gibellini D. et al. *Selective up-regulation of functional CXCR4 expression in erythroid cells by HIV-1 Tat protein*. Clin. Exp. Imuunol. 131: 428-435, 2003.
18. Re M.C. et al. *Antibody to HIV-1 Tat protein, a key molecule in HIV-1 pathogenesis. A brief review*. New. Microbiol. 2482: 197- 205, 2001.
19. Lewer A. *Regolatory protein of HIV*. Medical Virology 1: 155-163, 1991.

20. Emerman M. *HIV-1, Vpr and the cell cycle*. Curr Biol. 6(9):1096-1103, 1996.
21. Camaur D., Trono D. *Characterization of human immunodeficiency virus type 1 Vif particle incorporation*. J Virol. 70(9):6106-6111, 1996.
22. Cohen E.A., Subbramanian R.A., Gottlinger H.G. *Role of auxiliary proteins in retroviral morphogenesis*. Curr. Top. Microbiol. Immunol. 214: 219-235, 1996.
23. Pantaleo G., Graziosi C., Demaresi J.F. *HIV infection is active and progressive in lymphoid tissue during the clinically latent stage of disease*. Nature 362: 355-359,1993.
24. Romito A. et al. *Malignant neoplasm and AIDS*. Recenti Prog. Med. 88: 348-355,1997.
25. Del Canto M.C. *Mechanism of HIV infection of the central nervous system and pathogenesis of AIDS-dementia complex*. Neuroimaging Dlin. N. Am. 7: 231-241, 1997.
26. Re M.C. *Il laboratorio virologico nella diagnosi di infezione da HIV e nel monitoraggio del paziente infetto* (disponibile su: <http://www.bioetica-vssp.it/temi.html>).
27. Cassol S.A. et al. *Diagnosis of vertical HIV-1 transmission using the polymerase chain reaction and dried blood spot specimens*. J. Acquir. Immune. Defic.Syndr. 5 (2) : 113-119, 1992.
28. Guss, D. A. *The acquired immune deficiency syndrome: an overview for the emergency physician, 1994, Part 1*. J. Emerg. Med. 12, 375-384.

29. Morgan, D. et al. (2002) *Progression to symptomatic disease in people infected with HIV-1 in rural Uganda: prospective cohort study*. BMJ 324, 193-196.
30. Puchhammer-Stockl E., et al. *Low proportion of recent HIV infections among newly diagnosed cases of Hiv infection as shown by the presence of HIV-specific antibodies of low avidity*. J. of Clin. Microbiol.1: 497-498, 2005.
31. Van Binsbergen J. et al. *Strongly enhanced sensitivity of a direct anti-HIV-1/2 assay in seroconversion by incorporation of HIV p24 Ag detection: a new generation Vironostika HIV Uni-Form II*. J Virol.Methods, 76: 59-71, 1998.
32. Hirsch M.S. et al. *Antiretroviral drug resistance testing in adult HIV-1 infection*. Recommendations of an International AIDS Society-USA panel. JAMA 283: 2417-2426, 2000.
33. Hirsch M.S. et al. *Antiretroviral drug resistance testing in adult HIV-1 infection: implication for clinical management*. JAMA 279: 1984-1991,1998.
34. Roberts JD., Bebeneck K., Kukel T.A. *The accuracy of reverse transcriptase from HIV-1*. Science 242: 1171-1173, 1988.
35. Pillay D., Taylor S., Richman D.D. *Incidence and impact of resistance against approved antiretroviral drugs*. Rev. Med. Virol. 10: 231-253, 2000.

36. Re M.C. et al. *Analysis of HIV-1 drug resistance mutations by line probe assay and direct sequencing in a cohort of therapy naïve HIV-1 infected Italian patients.* BMC. Microbiol. 1 (1) 30, 2001.
37. Little S.J. et al. *Reduced antiretroviral drug susceptibility among patients with HIV infection.* JAMA 282 (12): 1142-1149, 1999.
38. Re M.C. et al. *Development of drug resistance in HIV-1 patients receiving a combination of stavudine, lamiduvine and efavirenz.* I. J. of Antimicrob. 20; 223-226, 2002.
39. Wegner S.A. et al. *Prevalence of genotyping and phenotyping resistance to anti-retroviral drugs in a cohort of therapy-naïve HIV-1 infected US military personnel.* AIDS 14, 1009-1015, 2000.
40. Flexner C. *HIV genotype and phenotype arresting resistance?* JAMA 283: 2442-2444, 2000.
41. Chiu, T.K. & Davies, D.R. *Structure and function of HIV-1 integrase,* *Curr. Top. Med. Chem.* 2004, 4, 965-977.
42. Chen, J.C. et al. *Crystal structure of the HIV-1 integrase catalytic core and Cterminal domains: a model for viral DNA binding,* *Proc. Natl. Acad. Sci. USA* 2000, 97, 8233-8238.
43. Wang J.Y., et al. *Structure a two domain fragment of HIV-1 integrase: implication for domain organization in the intact protein,* *EMBO J.* 2001, 20, 7333-7343.
44. Dyda, F. et al *Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases,* *Science* 1994, 266, 1981-1986.

45. Rice, P.A. & Baker, T.A. *Comparative architecture of transposases and integrase complex*. Nature Struct. Biol. 2001, 8, 302- 307.
46. Bujoucz, G. et al. *Binding of different divalent cations to the active site of avian sarcoma virus integrase and their effects on enzymatic activity*, J. Biol. Chem. 1997, 272, 18161-18168.
47. Young, W. & Steitz, T.A. *Recombining the structures of HIV integrase RuvC and RnaseH*. Structure 1995, 3, 131-134.
48. Beese, L.S. & Steitz, T.A. *Structural basis for the 3'-5' exonuclease activity of Escherichia Coli DNA polimerase I: a two metal ion mechanism*, EMBO J. 1991, 10, 25-33.
49. Grobler, J.A. et al. *Diketo acid inhibitor mechanism and HIV-1 integrase: implication for metal binding in the active site of phosphotransferases enzymes*, Proc. Natl. Acad. Sci. USA 2002, 99, 6661-6666.
50. Marchand, C. et al. *Metal-dependent inhibitor of HIV-1 integrase by β -diketo acids mutant (F185K/C280S)*, Mol.Pharmacol. 2003, 64, 600-609.
51. Chow, S.A. et al. *Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus*, Science 1992, 255, 723-726.
52. Chow, S.A. et al. *Juxtaposition of two viral DNA ends in a bimolecular disintegration reaction mediated by multimers of human immunodeficiency virus type 1 or murine leukemia virus integrase*, J. Virol. 1994, 68, 7869-7878.
53. Engelman, A. et al. *Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complexes* EMBO J. 1993, 12, 3269-3275.

54. Engelman, A. et al. *Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro.* J. Virol. 1992, 66, 6362-6369.
55. Puras-Lutzke, R.A. et al. *Characterization of the minimal DNA binding domain of the HIV integrase protein.* Nucleic Acids Res. 1994, 22, 4125-4131.
56. Ellison, V. and Brown, P.O. *A stable complex between integrase and viral DNA ends mediates human immunodeficiency virus integration in vitro.* Proc. Natl Acad. Sci. USA 1994, 91, 7316-7320.
57. Gerton, J.L. et al. *Effects of mutations in residues near the active site of human immunodeficiency virus type 1 integrase on specific enzyme-substrate interactions,* J. Virol. 1998, 72, 5046-5055.
58. Esposito, D. & Craigie, R. *Sequence specificity of viral end DNA binding by HIV-1 integrase reveals critical regions for protein-DNA interaction,* EMBO J. 1998 17, 5832-5843.
59. Bor, Y.C. *In-vitro integration of human immunodeficiency- virus type-1 cDNA into targets containing protein-induced bends.* Proc. Natl. Acad. Sci USA 1995, 92, 10334-10338.
60. Ferrin, T.E., Huang, C.C., Jarvis, L.E. & Langridge, R. *J. Mol. Graphics* 1988, 6, 13-27.
61. Goldgur, Y. et al. *Structure of HIV-1 integrase catalytic domain complexed with an inhibitor: a platform for antiviral drug design.* Proc. Natl. Acad. Sci. USA 1999, 96, 13040-13043.

62. Dayam, R. et al *Small-molecule HIV-1 integrase inhibitors: the 2001-2002 update*, Curr. Pharm. Design 2003, 9, 1789-1802.
63. Hazuda, D. et al. *Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells*. Science 2000, 287, 646-650.
64. Dayman. R et al. *Small-molecule HIV-1 integrase inhibitors: the 2001-2002 update*. Curr. Pharm. Design. 2003, 9, 1789-1802.
65. Pommier, Y. et al. *HIV-1 integrase as a target for antiviral drugs*. Ant. Chem. Chemoter. 1997, 8 (6), 463-482.
66. Neamati, N. et al. *Design and discovery of HIV-1 integrase inhibitors*. DDT 1997, 2, 487 498.
67. Costi, R. et al. *2,6-Bis-(3,4,5-trihydroxybenzylidene) derivatives of cyclohexanone: novel potent HIV-1 integrase inhibitors that prevent HIV-1 multiplication in cell-based assays*. Bioorg. Med. Chem. 2004, 12, 199-215.
68. Bukrinsky, M. et al, *Active nuclear import of human immunodeficiency virus type 1 preintegration complexes*. Proc. Natl.Acad. Sci. USA 1992, 89, 6580-6584.
69. Yoder, K. E. et al, *Repair of gaps in retroviral integration intermediates*. J. Virol. 2000, 74, 11191-11200.
70. Daniel, R. et al. *Evidence that stable retroviral transduction and cell survival following DNA integration depend on components of the homologous end joining repair pathway*. J. Virol. 2004, 78, 8573-8581.

71. Carteau, S. et al. *Chromosome structure and human immunodeficiency virus type 1 cDNA integration: centromeric alphoid repeats are disfavored target.* J. Virol. 1998, 72, 4005-4014.
72. Schoreder, A. R. et al. *HIV-1 integration in the human genome favors active genes and local hotspots.* Cell 2002, 110, 521- 529.
73. Bor, Y. C. et al. *Target-sequence preferences of HIV-1 integration complexes in vitro.* Virology 1996, 222, 283-288.
74. Pommier, Y. et al. *Integrase inhibitors to treat HIV/AIDS.* Nature Rew. 2005, 4, 236-248.
75. Di Santo, R. et al. *Novel bifunctional quinolonyl diketo acid derivatives as HIV-1 integrase inhibitors: design, synthesis, biological activities, and mechanism of action.* J. Med. Chem. 2006, 49, 1939-1945.
76. Goldgur, Y. et al, *Three new structure of the core domain of HIV-1 integrase: an active site that binds magnesium.* Proc. Natl. Acad. Sci. USA 1998, 95, 9150-9154.
77. Goldgur, Y. et al. *Structure of the HIV-1 integrase catalytic domain complexed with an inhibitor: A platform for antiviral drug design.* Proc. Natl. Acad. Sci. USA 1999, 96, 13040- 13043.
78. Lin, J.-H. et al. *Computational drug design accommodating receptor flexibility: the relaxed complex scheme.* J. Am. Chem. Soc. 2002, 124, 5632-5633.
79. Morris, G. M. et al. *Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function.* J. Comput. Chem. 1998, 19, 1639-1662.

80. Greenwald, J. et al. *The mobility of an HIV-1 integrase active site loop is correlated with catalytic activity*. Biochem. 1999, 38, 8892-8898.
81. Lins, R. D. et al, *Molecular dynamics studies on the HIV-1 integrase catalytic domain*. Biophys. J. 1999, 76, 2999-3011.
82. Podtelezhnikov, A. A. et al. *Modeling HIV-1 integrase complexes based on their hydrodynamic properties*. Biopolymers 2003, 68, 110-120.
83. Van Gent, D. C. et al. *Mutational analysis of the integrase protein of human immunodeficiency virus type 2*. Proc. Natl. Acad. Sci. USA 1992, 20, 9598-9602.
84. Marchand, C. et al, *Structural determinant for HIV-1 integrase inhibition by α -diketo acids*. J. Biol. Chem. 2002, 277, 12596-12603.
85. Van Gent, D. C. et al. *Mutational analysis of the integrase protein of human immunodeficiency virus type 2*. Proc. Natl. Acad. Sci. USA 1992, 20, 9598-9602.
86. Jenkins, T. M. et al, *Critical contacts between HIV-1 integrase and viral DNA identified by structure-based analysis and photo-crosslinking*. EMBO J. 1997, 16, 6849-6859.
87. Maignan, S. et al. *Crystal structures of the catalytic domain of HIV-1 integrase free and complexed with its metal cofactor: high level of similarity of the active site with other viral integrases*. J. Mol. Biol. 1998, 282, 359-368.
88. Bujacz, G. et al. M. *The catalytic domain of avian sarcoma virus integrase: conformation of the active-site residues in the presence of divalent cations*. Structure 1996, 4, 89-96.

89. De Luca, L. et al. *Analysis of the full-length integrase-DNA complex by a modified approach for DNA docking*. *Biochem Biophys Res Commun.* 2003, 310, 1083-1088.
90. Adesokan, A.A. et al. *Prediction of HIV-1 integrase/viral DNA interactions in the catalytic domain by fast molecular docking*. *J. Med. Chem.* 2004, 47, 821-828.
91. Heuer, T.S. and Brown, P.O. *Photo-cross-linking studies suggest a model for the architecture of an active human immunodeficiency virus type 1 integrase-DNA complex*. *Biochemistry*, 1998, 37, 6667-6678.
92. Davies, D.R et al. *Three-dimensional structure of the Tn5 synaptic complex transposition intermediate*. *Science*, 2000, 289, 77-85.
93. Rice, P. A.; Baker, T. A. *Comparative architecture of transposase and integrase complexes*. *Nat. Struct. Biol.* 2001, 8, 302-307.
94. Ason, B. et al. *Targeting Tn5 transposase identifies human immunodeficiency virus type 1 inhibitors*. *Antimicrob. Agents Chemother.* 2005, 49, 2035-2043.
95. Barreca, M.L. et al. *Binding mode prediction of strand transfer HIV-1 integrase inhibitors using Tn5 transposase as a plausible surrogate model for HIV-1 integrase*. *J. Med. Chem.* 2006, 49, 3994-3997
96. Wielens, J.; Crosby, I.T.; Chalmers, D.K. *A three-dimensional model of the human immunodeficiency virus type 1 integration complex*. *J. Comput. Aided Mol. Des.* 2005, 19, 301-317.
97. Neamati, N.; et al. *Metal-dependent inhibition of HIV-1 integrase*. *J. Med. Chem.* 2002, 45, 5661-5670.

98. Hazuda, D.J. et al. *A naphthyridine carboxamide provides evidence for discordant resistance between mechanistically identical inhibitors of HIV-1 integrase*. Proc Natl Acad Sci U S A. 2004, 101, 11233-11238.
99. Allison, A. A. et al. *Probing HIV-1 integrase inhibitor binding sites with position-specific integrase-DNA cross-linking assays*. Mol. Pharmacol. 2007, 71, 893-901.
100. Jenkins, T. M. et al. *Critical contacts between HIV-1 integrase and viral DNA identified by structure-based analysis and photo-crosslinking*. EMBO J. 1997, 16, 6849-6859.
101. SYBYL Molecular Modeling System, TRIPOS Assoc., St. Louis, MO.
102. Case, D. A. et al. *AMBER 9*, University of California, San Francisco, 2004.
103. Cornell, W. D. et al. *A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules*. J. Am. Chem. Soc. 1995, 117, 5179-5197.
104. Jorgensen, W.L. et al. *Comparison of simple potential functions for simulating liquid water*. J. Chem. Phys. 1983, 79, 926-935
105. Ryckaert, J. P. et al. *Numerical integration of the Cartesian equations of motion for a system with constraints: Molecular dynamics of n-alkanes*. J. Comput. Phys. 1977, 23, 327-333.
106. Huey, R.; et al. . *A semiempirical free energy force field with charge-based desolvation*. J. Comput. Chem. 2007, 28, 1145-1152.
107. Head, J.; Zerner, M. C. *A Broyden-Fletcher-Goldfarb-Shanno optimization procedure for molecular geometries*. Chem. Phys. Lett. 1985, 122, 264-274.

108. Bor, Y.C. *In-vitro* integration of human immunodeficiency- virus type-1 *cDNA* into targets containing protein-induced bends. Proc. Natl. Acad. Sci USA 1995, 92, 10334-10338.

Siti Internet consultati

- www.avert.org/virus.htm
- www.medicinenet.com
- www.aidsinfo.nih.gov
- www.aids.med.unibo.it
- www.thebody.com
- www.cdc.gov/hiv
- www.ministerosalute.it/aids
- www.roche-applied-science.com/lightcyclr
- www.ncbi.nlm.nih.gov
- www.aids.org

Ringraziamenti

Ringrazio i Proff. Ettore Novellino ed Antonio Lavecchia che attraverso il loro supporto morale e materiale mi hanno permesso di condurre e portare a termine questo mio progetto di dottorato attraverso l'applicazione di metodiche computazionali.

Ringrazio il Prof. R. Di Santo ed il suo gruppo di ricerca, dell'Istituto Pasteur-Fondazione Cenci Bolognetti, dell'Universita' di Roma "La Sapienza" che si sono occupati della parte sintetica del progetto.

Ringrazio i Proff. Y.Pommier e C. Marchand ed il loro gruppo di ricerca del National Institutes of Health di Bethesda, che si sono occupati di eseguire i saggi di attività biologica e citotossicità del progetto.