

# IDENTIFICATION D'UN NOUVEL INTERACTANT SPECIFIQUE DE L'INTEGRASE DU RETROVIRUS ENDOGENE PORCIN : LA PROTEINE CELLULAIRE BRD2

Kathy Gallay

▶ **To cite this version:**

Kathy Gallay. IDENTIFICATION D'UN NOUVEL INTERACTANT SPECIFIQUE DE L'INTEGRASE DU RETROVIRUS ENDOGENE PORCIN : LA PROTEINE CELLULAIRE BRD2. Virologie. 2012. hal-01470739

**HAL Id: hal-01470739**
**https://ephe.hal.science/hal-01470739**

Submitted on 17 Feb 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE

ECOLE PRATIQUE DES HAUTES ETUDES

Sciences de la Vie et de la Terre

MEMOIRE

Présenté par

Kathy GALLAY

pour l'obtention du diplôme de l'Ecole Pratique des Hautes Etudes

---

# IDENTIFICATION D'UN NOUVEL INTERACTANT SPECIFIQUE DE L'INTEGRASE DU RETROVIRUS ENDOGENE PORCIN : LA PROTEINE CELLULAIRE BRD2

---

Soutenu le 27 juin 2012, devant le jury suivant :

| | |
|---|---|
| Pr Thierry DUPRESSOIR | Président du jury |
| Dr Corinne RONFORT | Tuteur scientifique |
| Pr Christophe TERZIAN | Tuteur pédagogique |
| Dr Marc LAVIGNE | Rapporteur |
| Dr Yannick BLANCHARD | Examinateur |

Mémoire préparé sous la direction de :

Dr Corinne RONFORT
UMR 754 INRA-ENVL-UCBL-EPHE
« Rétrovirus et pathologie comparée »
SFR128 Biosciences Lyon Gerland
Equipe : *Rétrovirus, Intégration Virale et Vaccins*
corinne.ronfort@univ-lyon1.fr

Directeur : Pr Jean-François MORNEX

Et de

Pr Christophe TERZIAN
UMR 754 INRA-ENVL-UCBL-EPHE
« Rétrovirus et pathologie comparée »
SFR128 Biosciences Lyon Gerland
Equipe : *Infection et Rétrovirus Endogènes*
christophe.terzian@univ-lyon1.fr

ECOLE PRATIQUE DES HAUTES ETUDES
SCIENCE DE LA VIE ET DE LA TERRE

# IDENTIFICATION D'UN NOUVEL INTERACTANT SPECIFIQUE DE L'INTEGRASE DU RETROVIRUS ENDOGENE PORCIN : LA PROTEINE CELLULAIRE BRD2

Kathy GALLAY

## Résumé

Les rétrovirus sont des virus à ARN enveloppés dont le génome est converti en un ADN double brin au cours du cycle viral. Cet ADNc est inséré dans le génome de la cellule infectée par une enzyme virale : l'intégrase (IN). De nombreuses études ont montré le rôle auxiliaire joué par des protéines cellulaires durant l'étape d'intégration rétrovirale, dont certains peuvent favoriser le processus intégratif. Récemment, une protéine cellulaire nommée Brd2 (pour BRomoDomain containing protein 2) a été identifiée en interaction avec les intégrases du virus HIV-1 (Human Immunodeficiency virus type1) et celle du virus MLV (Murine Leukemia virus). La protéine Brd2 est un régulateur transcriptionnel interagissant avec certaines formes d'histones acétylés présents au niveau de la chromatine. Notre étude a pour objectif de comprendre le rôle de la protéine Brd2 au cours du cycle viral et notamment lors du processus intégratif.

Ce travail est réalisé sur des modèles viraux représentés par des rétrovirus aviaires, humains, murins et porcins. Nous avons ainsi pu mettre en évidence une interaction entre l'IN du rétrovirus endogène porcin et la protéine Brd2 humaine par Far Western blot et par colocalisation cellulaire. Nous avons aussi caractérisé les domaines d'interaction entre ces deux partenaires. Nous cherchons actuellement à comprendre l'importance de ce facteur cellulaire, son rôle sur les activités catalytiques de l'intégrase et sur la réplication virale du PERV (Porcine Endogenous RetroVirus).

Les résultats obtenus auront un impact sur les connaissances fondamentales des interactions hôte-rétrovirus et sur la définition de cibles potentielles dans le cadre de la lutte contre les rétrovirus.

Mots clés : Rétrovirus, Intégrase, HIV-1, PERV, LEDGF/p75, Brd2

# SOMMAIRE

Liste des abréviations principales
**Analyse bibliographique**
*Préambule* ..... *1*
*Chapitre 1 : Présentation des rétrovirus* ..... *2*
1 Les rétrovirus ..... *2*
1.1 Les orthoretrovirinae ..... 2
1.2 Les spumaretrovirinae ..... 3
1.3 Les rétrovirus endogènes ..... 3
*2 Structure et organisation génomique des rétrovirus* ..... *4*
2.1 Structure de la particule virale ..... 4
2.2 Organisation génomique des rétrovirus ..... 4
2.2.1 Structure du génome viral : ..... 4
2.2.2 LôARN génomique et les ARN messagers viraux ..... 5
2.2.3 Les protéines virales enzymatiques et de structure ..... 5
2.2.4 Les protéines accessoires et régulatrices ..... 6
*3 Cycle de réplication des rétrovirus* ..... *7*
3.1 La phase précoce ..... 7
3.2 La phase tardive ..... 8
*Chapitre 2 : Lôintégration rétrovirale, l'integrase virale et le co-facteur LEDGF/P75* ..... *9*
*1 L'intégration rétrovirale* ..... *9*
1.1 Composition et import nucléaire du complexe de pré-intégration ..... 9
1.1.1 Composition des complexes pré-intégratifs ..... 9
1.1.2 Transport des CPI ..... 9
1.2 Lôintégration virale ..... 10
1.2.1 Mécanisme ..... 10
1.2.2 Sites dôintégration ..... 11
1.2.3 Les formes virales circulaires ..... 12
*2 La protéine intégrase* ..... *12*
2.1 Structure en domaines de lôintégrase ..... 12
2.1.1 Le domaine N-terminal ..... 12
2.1.2 Le domaine central ..... 12
2.1.3 Le domaine C-terminal ..... 13
2.2 Structure oligomérique de lôintégrase ..... 13
*3 LEDGF/P75 : co-facteur cellulaire de lôintégrase du HIV-1* ..... *14*
3.1 Présentation et organisation génomique ..... 14
3.2 Fonctions ..... 15
3.3 LEDGF/p75 et Intégrase ..... 15
3.4 LEDGF/p75 dans le cycle viral ..... 16
3.5 Cas de LEDGF/p52 ..... 18
3.6 Conclusion ..... 18
*Chapitre 3 : Les protéines à bromodomaines et Brd2* ..... *19*
*1 Généralités et remodelage de la chromatine* ..... *19*
*2 Les protéines à bromodomaines* ..... *21*
*3 Les protéines BET* ..... *21*
*4 La protéine Brd2 humaine* ..... *21*
4.1 Fonctions de Brd2 : ..... 22

4.2 Localisation cellulaire de Brd2. ....23
*5 Les protéines BET en interaction avec d'autres virus ....23*
5.1 Brd4 et Papillomavirus :....23
5.2 Protéines BET et KSHV :....24
5.3 Protéines BET et autres herpes virus : ....24
5.4 Protéines BET et rétrovirus :....24
5.5 Conclusion....25
5.6 Brd2 et Intégrase ....25

# Liste des abréviations principales

| | |
|---|---|
| ADN | Acide désoxyribonucléique |
| ADNc | ADN complémentaire |
| ALSV | Avian Leukemia and Sarcoma Virus |
| APS | Ammonium persulfate |
| ARN | Acide ribonucléique |
| Att | Attachement |
| BD | Bromodomaine |
| BET | Bromodomain et extraterminal domain |
| BRD2 | Bromodomain containing protein 2 |
| CCD | Core Catalytique |
| CPI | Complexe de préintégration |
| CTD | Domaine C-terminal |
| DMSO | Dimethyl sulfoxyde |
| DO | Densité optique |
| DTT | Dithiotréïtol |
| EDTA | Ethylenediaminetetraacetic acid |
| ET | Extra terminal domain |
| GFP | Green Fluorescente protein |
| HAT | Histone acétyltransférase |
| hBrd2 | Brd2 humaine |
| HRP | hepatoma-derived growth factor-related protein |
| IBD | Integrase binding domain |
| IN | Integrase |
| IPTG | Isopropyl-b-D-thiogalactopyranoside |
| kDa | Kilo Dalton |
| LEDGF/p75 | Lens Epithelium-Derived Growth Factor |
| LTR | Long terminal repeat |
| M | Molaire (mol.l-1) |
| MLV | Murine Leukemia Virus |
| MLV | Murine Leukemia Virus |
| Mo-MLV | Moloney Murine Leukemia Virus |
| NLS | Signal de localisation nucléaire |
| nt | Nucléotide |
| NTD | Domaine N-terminal |
| pb | Paire de base |
| PCR | Polymerase Chain Reaction |
| PERV | Porcine endogenous retrovirus |
| PFV | Prototype foamy virus |
| RAV | Rous Associated Virus |
| RSV | Rous Sarcoma Virus |
| RT | Reverse transcriptase |
| RTC | Reverse transcriptase complexe |
| SDS | Sodium Dodecyl Sulfate |
| SIDA | Syndrome de lôImmunodéficience Acquise |

| | |
|---|---|
| TEMED | N,N,N',N'-Tetramethylethylenediamine |
| VIH-1 | Virus de lôImmunodéficience Humaine de type 1 |

# Préambule

Dans le cadre de mon diplôme de l'Ecole Pratique des Hautes Etudes, j'ai pu intégrer un projet de recherche portant sur les partenaires cellulaires de la protéine intégrase des rétrovirus. Ce projet a été réalisé au sein de l'équipe « Rétrovirus, Intégration Virale et Vaccins » dirigée par le Dr Corinne Ronfort.

Les rétrovirus sont des virus à ARN, enveloppés, trouvés dans de nombreuses espèces animales de rente (oiseaux, ovins, caprins, bovins, équins…), les animaux de compagnie (chat) et l'homme. Le représentant le plus connu est le Virus de l'Immunodéficience Humaine (HIV) responsable du Syndrome de l'ImmunoDéficience Acquise (SIDA). Le cycle réplicatif des rétrovirus est caractéristique : le génome ARN est converti en une molécule d'ADN double brin puis celle-ci est insérée dans le génome de la cellule infectée par une enzyme virale appelée intégrase (IN).

Un des axes de recherche de l'équipe porte sur l'identification et la caractérisation d'interactants cellulaires de l'IN. De nombreuses études ont montré l'importance des protéines cellulaires durant l'étape d'intégration rétrovirale. Certaines peuvent favoriser le processus intégratif.

Dans l'introduction de ce document, nous présenterons les caractéristiques majeures des rétrovirus, à savoir leur classification, leur structure, leur organisation génomique et leur cycle réplicatif. Puis, nous étudierons plus particulièrement l'étape d'intégration rétrovirale et la protéine IN ainsi qu'un partenaire cellulaire de l'IN des lentivirus très bien décrit ces dernières années : la protéine LEDGF/p75. Enfin, nous nous intéresserons aux protéines de la famille des bromodomaines et extraterminal domaine (BET) et plus particulièrement à la protéine Brd2 qui a, récemment, été identifiée en interaction avec l'IN d'un gammarétrovirus. A l'issue de cette analyse bibliographique, nous présenterons les objectifs de l'étude.

Nous présenterons ensuite les méthodes développées dans ce travail puis les résultats acquis. Ceux-ci seront enfin discutés dans une dernière partie.

.

# Chapitre 1 : Présentation des rétrovirus

## 1 Les rétrovirus

Les R*etroviridae* ou rétrovirus représentent une famille large et diversifiée de virus définie par des critères taxonomiques communs incluant la structure et la composition du virion ainsi que ses propriétés réplicatives. La famille des R*étroviridae* se subdivise en deux sous familles : les *orthoretrovirinae* et les *spumaretrovirinae*. Ces deux sous familles comprennent au total 7 genres définis par lôInternational Committee on Taxonomy of Viruses (ICTV).

### 1.1 Les orthoretrovirinae

Les *orthoretrovirinae* se divisent en six genres :

- Le genre **Alpharétrovirus** est constitué de rétrovirus infectant les oiseaux, tels les virus de la famille des Virus de la Leucose Aviaire (ALV) (comme le RAV-1 par exemple) ou de la famille des virus du Sarcome Aviaire (ASV) (comme le Virus du Sarcome de Rous (RSV)). Les ASV sont caractérisés par la présence dans leur génome de séquences oncogéniques et sont responsables de tumeurs aigues avec des périodes de latence très courtes. Ceux-ci sont défectifs pour la réplication et ont besoin de la présence dôun virus compétent pour être infectieux. Les ALV ne possèdent pas dôoncogène mais sont activateurs dôoncogènes cellulaires (Weiss, 2006). Ces deux types viraux ASV et ALV sont regroupés au sein dôune même espèce : les virus ASLV.
- Le genre **Bétarétrovirus** dont le représentant majeur est le Virus de la Tumeur Mammaire de la souris ou MMTV. Ces rétrovirus ne possèdent pas de séquences oncogéniques.
- Le genre **Gammarétrovirus** est composé de virus infectant une grande variété de mammifères. Il comprend des virus simples, compétents ou défectifs et certains présentent des séquences oncogéniques comme le virus de la leucémie féline (FeLV). On y trouve aussi le virus de la leucémie murine (MLV) et le virus endogène porcin (PERV)
- Le genre **Deltarétrovirus** est constitué de virus infectant les mammifères et notamment lôhomme. Ils sont responsables de leucémies. Ils sont représentés par les virus de la Leucose Bovine (BLV) et des Virus de la Leucémie des cellules T Humaines 1 et 2 (HTLV-1 et HTLV-2).
- Le genre **Epsilonrétrovirus** est composé de virus qui infectent uniquement les poissons comme le Virus du Sarcome Dermique du saumon Walleye (WDSV)

- Le genre **Lentivirus** est constitué de rétrovirus non oncogènes mais cytopathogènes. Ces virus ont été isolés chez différentes espèces animales dont les primates (HIV-1, HIV-2, SIV), les félins (FIV) ou encore les bovins (BIV). Ils sont responsables de maladies chroniques à évolution lente dôoù lôappellation « Lentivirus » signifiant virus lent en latin, et conduisent à des immunodéficiences à l'origine de diverses maladies opportunistes. Après infection, la période dôincubation peut sôétendre sur des mois voire des années avant que les signes cliniques ne se manifestent. Les lentivirus peuvent être divisés en deux groupes sur la base de leur tropisme cellulaire : le Virus de lôAnémie Infectieuse Equine (EIAV), le Virus de lôArthrite et de lôEncéphalite Caprine (CAEV) et le Virus Visna-Maedi (MVV) infectent la lignée monocyte-macrophage, alors que HIV, SIV et FIV infectent, en plus des macrophages, les lymphocytes T portant lôantigène de surface de la classe de différenciation 4 (CD4+). Le représentant le plus connu de ce genre est le HIV chez lôhomme, lôagent causal du SIDA.

## 1.2 Les spumaretrovirinae

Les *spumaretrovirinae* quant à eux ne se composent que du genre spumavirus. Ils comprennent les Virus Foamy Simien, Félin, Equin, Bovin (SFV, FFV, EFV, BFV). Ces virus sont caractérisés par un fort effet cytopathogène en culture et sont connus pour causer des infections persistantes sans signe clinique.

## 1.3 Les rétrovirus endogènes

Certains rétrovirus sont endogènes. Ces rétrovirus endogènes ou ERV sont des virus ayant intégré leur génome dans la lignée germinale dôun hôte ; ils se transmettent ensuite de façon verticale de génération en génération. Les rétrovirus endogènes humains (HERV) représentent 8% du génome humain.

**Cas du rétrovirus endogène porcin : PERV**

Le rétrovirus endogène porcin ou PERV appartient au genre gammarétrovirus et est phylogénétiquement lié aux virus MLV, FeLV, GaLV. Ce rétrovirus est intégré dans le génome de tous les porcs. Selon leur tropisme, on distingue 3 sous-groupes infectieux : PERV-A et PERV-B infectent les cellules humaines et porcines *in vitro* et PERV-C infecte exclusivement les cellules porcines (Denner, 2008). La plupart des provirus sont défectifs et incapables de produire des virus compétents pour la réplication. Dans des cellules porcines, le PERV-C est capable de recombiner avec la séquence de lôenveloppe virale du PERV-A pour produire un recombinant appelé PERV A/C. Ce recombinant est capable dôinfecter les cellules humaines avec une plus grande efficacité.

# 2 Structure et organisation génomique des rétrovirus

## 2.1 Structure de la particule virale

Les particules virales ou virions ont un diamètre de 80 à 120 nanomètres et la taille du génome des rétrovirus varie de 7 à 12 kb. Ce génome est composé de deux molécules dôARN simple brin, de polarité positive fortement condensées par association avec la protéine de nucléocapside (NC) qui protège le génome. Dans le complexe ARN/NC, on retrouve aussi des protéines enzymatiques virales : la transcriptase inverse (RT) et lôintégrase (IN). Le tout est protégé par une structure protéique en forme de cône tronqué appelée capside (CA). Cette dernière est entourée dôune matrice protéique (MA) qui tapisse la face interne de lôenveloppe virale et contient une autre enzyme virale : la protéase (PR). Sur lôenveloppe phospholipidique externe dôorigine cellulaire sont hérissées deux sortes de glycoprotéines : (i) la glycoprotéine transmembranaire (TM) qui traverse la bicouche lipidique. Cette protéine est responsable de la fusion de lôenveloppe virale avec la membrane cellulaire. (ii) La glycoprotéine de surface (SU) qui occupe une position plus périphérique et assure la reconnaissance du virus sur des récepteurs spécifiques de la cellule cible lors de lôinfection. La particule virale contient aussi des molécules cellulaires encapsidées telles que lôARNt, la cyclophiline A (Freed, 2001).

## 2.2 Organisation génomique des rétrovirus

### *2.2.1 Structure du génome viral :*

Selon la composition du génome, on distingue deux groupes :

**Les rétrovirus à génome simple** : Leur génome est composé de 3 gènes qui sont :

- le gène *gag* (pour *groupe-specific antigen*) qui code les protéines de structures de la particule virale (CA, MA, NC),
- le gène *pol* (pour polymerase) qui code les enzymes impliquées dans la réplication et lôintégration virale (IN, RT, PR),
- le gène *env* (pour enveloppe) qui code les glycoprotéines de lôenveloppe virale (TM, SU).

Les alpharétrovirus, des betarétrovirus, les epsilonrétrovirus et les gammarétrovirus sont des rétrovirus à génome simple.

**Les rétrovirus à génome complexe** : En plus des 3 gènes *gag, pol, env*, leur génome code des protéines additionnelles (protéines régulatrices ou accessoires). Les deltarétrovirus, les spumavirus et des lentivirus sont des rétrovirus à génome complexe.

A lôétat de provirus, le génome viral est encadré par deux régions non codantes, appelées LTR (pour Long Terminal Repeat) impliquées dans lôintégration virale et dans la régulation de lôexpression du génome viral. Ces séquences se divisent en trois éléments : une région unique nommée U3 suivie de la région R (Repeat) puis de la région U5. Le LTR du coté 5ô est un promoteur puissant de la transcription. Du coté 3ô, le LTR fournit le signal de coupure qui précède la polyadénylation.

### *2.2.2 LôARN génomique et les ARN messagers viraux*

Pour les rétrovirus simple (comme les virus de type ALV), lôexpression du provirus forme deux ARN messagers : un transcrit primaire non épissé appelé ARN génomique ou ARNm *gag-pol* et un transcrit mono-épissé appelé ARN sous génomique ou ARNm *env*. Pour le HIV-1 (rétrovirus complexe), il y a en plus, des transcrits multi-épissés permettant la traduction des protéines accessoires et régulatrices.

LôARN génomique débute par les régions R-U5 et se termine par les régions U3-R. En aval de la région U5 se trouve une séquence PBS (Primer Binding Site) qui sert de site de fixation pour un ARN de transfert à partir duquel sôamorce la transcription inverse. On trouve aussi une région Ψ correspondant au signal dôencapsidation de lôARN viral génomique non épissé dans les particules en formation. En amont de la région U3, on trouve une séquence riche en A/G appelée séquence PPT (polyPurine Tract) qui, lors de la transcription inverse, résiste à lôaction de la RNase H et sert dôamorce ARN pour la synthèse du second brin dôADN. Chez les lentivirus, cette séquence est aussi présente en position centrale dans le génome et est appelée central polypurine tract ou cPPT.

### *2.2.3 Les protéines virales enzymatiques et de structure*

A partir des deux ARN messagers, trois précurseurs polyprotéiques (Pr) sont produits : Pr Gag, Pr Gag-Pol et Pr Env.

LôARNm *env* permet la traduction dôun précurseur polypeptidique Pr Env qui, après glycosylation, est clivé par une protéase cellulaire pour donner les glycoprotéines de lôenveloppe virale : la TM et la SU.

LôARNm *gag-pol* permet la traduction de deux polyprotéines : le Pr Gag et le Pr Gag-Pol : En 3ô du gène *gag* se trouve un codon stop qui est, dans 90% des cas, interprété correctement par le ribosome et permet la synthèse dôun précurseur polypeptidique Pr Gag. Lôenvironnement du

codon stop permet, dans 10 % des cas, un décalage du cadre de lecture et la synthèse dôun précurseur polypeptidique Pr Gag-Pol.

Le précurseur polypeptidique Pr-Gag, une fois clivé par la protéase virale (PR), donne trois protéines majeures qui sont la capside (CA), la matrice (MA) et enfin, la nucléocapside (NC).

Le précurseur polypeptidique Pr-Gag-Pol permet lôobtention des différentes enzymes virales après clivage du Pr-Gag-Pol par la protéase virale. Ces enzymes sont : la transcriptase inverse (RT) et lôintégrase (IN). La protéase (PR) est, quant à elle, codée soit par le gène *gag* soit par le gène *pol* selon le type de rétrovirus (*gag* pour les ALV et *pol* pour HIV). La RT est une ADN polymérase ARN dépendante qui permet la transcription inverse de lôARN viral en un double brin dôADN. Elle possède aussi une activité RNase H permettant dôhydrolyser le brin dôARN dôun hybride ADN-ARN constitué au cours du processus de transcription inverse. LôIN a pour rôle majeur dôintégrer lôADN double brin dans le génome de la cellule infectée. Elle possède aussi dôautres fonctions qui seront détaillées dans le chapitre 2. Enfin, la PR clive les précurseurs polypeptidiques au niveau de sites précis conduisant à lôobtention de protéines virales matures.

### *2.2.4 Les protéines accessoires et régulatrices*

Les rétrovirus complexes possèdent, en plus des gènes *gag*, *pol* et *env*, des gènes supplémentaires codant des protéines accessoires et régulatrices qui jouent un rôle important dans lôinteraction hôte-pathogène et vont avoir un impact significatif sur la réplication virale (Li et al., 2005). Dans le cas dôHIV-1 par exemple, le génome de ce rétrovirus comprend deux gènes codant des protéines régulatrices, Tat et Rev. La protéine Tat (Transactivator of transcription) est un transactivateur puissant de la transcription du HIV-1. La protéine Rev (Regulator of Expression of the Virion) favorise lôexport des ARNs messagers viraux non et mono-épissés vers le cytoplasme.

Le génome du HIV-1 contient aussi des séquences codant des protéines accessoires telles que Nef, Vif, Vpr, Vpu (Li et al., 2005) :

La protéine Nef (Negative Regulator Factor) possède quatre activités indépendantes au cours de la réplication. Elle affecte lôexpression de certaines protéines cellulaires de surface, interfère avec des signaux de transduction cellulaire, augmente lôinfectiosité du virion et de la réplication.

La protéine Vif (Virus Infectivity Factor) est essentielle pour la réplication du virus dans les lymphocytes du sang périphérique, les macrophages et certaines lignées comme les cellules non-permissives. Vif se fixe au facteur cellulaire APOBEC3G (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G). Lôenzyme cellulaire APOBEC3G constitue lôun des mécanismes de défense antivirale naturelle des cellules de mammifères contre les rétrovirus. APOBEC3G modifie la composition génétique du virus en désaminant lôARN viral ce qui forme

des virus non infectieux. Le rôle essentiel de la protéine Vif est de neutraliser cette ligne de défense en provoquant la dégradation dôAPOBEC3G, permettant ainsi la propagation du virus.

La protéine Vpr (Viral Protein R) possède de nombreuses fonctions pour la réplication du virus notamment dans lôimport nucléaire du complexe de pré-intégration et de lôapoptose, dans lôarrêt du cycle cellulaire en phase G2 et dans la transactivation des LTR viraux et de certains gènes cellulaires.

La protéine Vpu (Viral Protein U) augmente la libération des nouvelles particules virales et induit la dégradation des récepteurs CD4. Elle est capable de bloquer la Téthérine, une protéine cellulaire qui retient le virus à la surface de la cellule (Li et al. 2005).

# 3 Cycle de réplication des rétrovirus

Le cycle de réplication des rétrovirus, est divisé en deux parties : les phases dites précoces comprenant les étapes dôentrée du virion jusquôà lôintégration du provirus et les phases dites tardives qui comprend toutes les étapes qui ont lieu après lôintégration (Freed, 2001).

## 3.1 La phase précoce

**Attachement et entrée**

La première étape de lôinfection est lôattachement du virus à la surface de la cellule. La pénétration de la particule virale dans la cellule cible se fait par lôinteraction entre la glycoprotéine de surface SU et un récepteur spécifique présent à la surface de la cellule. Pour le HIV-1, le récepteur cellulaire est le CD4 présent à la surface des lymphocytes T4 mais cette interaction nôest pas suffisante pour permettre lôinternalisation du virus. Elle entraine cependant un changement de conformation de la glycoprotéine de surface SU qui va permettre son interaction avec un corécepteur transmembranaire (récepteurs aux chemokines) : CCR5 ou CXCR4. Cette seconde interaction va libérer le peptide fusion et lôenveloppe virale et la membrane cellulaire vont fusionner.

**La décapsidation**

La capside est alors délivrée dans le cytoplasme et se désagrège libérant ainsi un complexe nucléoprotéique appelé complexe de transcription inverse (reverse transcriptase complexe ou RTC) comprenant le génome, les protéines enzymatiques virales et différentes molécules dôorigine cellulaire.

**La transcription inverse**

Lôétape de transcription inverse est réalisée au sein du RTC. La transcriptase inverse virale ou rétrotranscriptase, va copier lôARN viral génomique en un ADN double brin. Cette étape conduit à la création de séquences particulières aux extrémités de l'ADN proviral : les LTRs composées des séquences U3 R U5, qui contiennent les promoteurs et terminateurs de la transcription.

**Import nucléaire et intégration**

Cet ADN double brin, associé à lôintégrase et à diverses protéines virales et cellulaires, forme un complexe de pré intégration (CPI). Celui-ci est importé dans le noyau de la cellule (pour les lentivirus) puis intégré au sein du génome de la cellule cible. Ces processus seront décrits dans le chapitre 2.

## 3.2 La phase tardive

**La transcription**

Le provirus dépend de lôARN polymérase cellulaire pour sa transcription. Les transcrits viraux ont donc la même structure que tout ARNm eucaryote, avec une coiffe à lôextrémité 5ô et une queue de polyadénines (poly A) à lôextrémité 3ô et ils sont épissés par les enzymes cellulaires.

**La traduction**

Les ARNm viraux sont traduits par la machinerie cellulaire permettant la production des éléments structurels du virion, des protéines de lôenveloppe ainsi que des enzymes virales.

**Assemblage et bourgeonnement**

Après la traduction, les protéines virales assemblées sôaccumulent sous la membrane plasmique pour former, avec lôARN viral, une particule virale. Cette particule virale bourgeonne à la surface de la cellule entrainant avec elle une partie de la bicouche lipidique de la membrane cellulaire.

**Maturation**

Les particules virales immatures ainsi libérées sont ensuite maturées par la protéase virale incorporée dans le virion. La protéase clive ensuite les différents précurseurs polyprotéiques permettant ainsi de former des virions infectieux.

# Chapitre 2 : Lôintégration rétrovirale, l'integrase virale et le co-facteur LEDGF/P75

Pour tous les rétrovirus, lôintégration de lôADN viral au sein du génome de la cellule hôte est une étape indispensable au cycle viral. Pour cela, lôADN viral doit être importé dans le noyau de la cellule infectée.

## 1 L'intégration rétrovirale

### 1.1 Composition et import nucléaire du complexe de pré-intégration

#### *1.1.1 Composition des complexes pré-intégratifs*

Après la transcription inverse, lôADN viral néosynthétisé reste associé à des protéines virales et cellulaires au sein dôun complexe nucléoprotéique de haut poids moléculaire. Ce complexe est appelé complexe de pré-intégration (CPI) (Suzuki and Craigie, 2007). Dans le CPI, lôIN, sous forme de tetramère, se fixe aux 2 LTRs de lôADN viral et recrute des cofacteurs viraux et cellulaires (Van Maele et al., 2006). La composition du CPI nôest pas encore bien définie mais des études ont révélé que le CPI de HIV-1 retient vraisemblablement plusieurs composants viraux du RTC tels que les protéines IN, MA, NC, RT et la protéine Vpr (Miller et al., 1997). De plus, plusieurs protéines cellulaires ont été retrouvées associées au CPI de HIV-1 comme les protéines HMG1 (High Mobility Group), BAF (barrier to auto integration factor), LAP2Ŭ (Lamina-associated polypeptide 2 Ŭ) et LEDGF/p75 (Lens epithelium growth factor) (Farnet and Bushman, 1997; Lin and Engelman, 2003; Llano et al., 2004a; Turlure et al., 2004; Van Maele et al., 2006).

#### *1.1.2 Transport des CPI*

Le transport vers le noyau et lôimport nucléaire sont mal connus. Lôadressage nucléaire du CPI est probablement réalisé grâce à son interaction avec le réseau de microtubules. Une étude utilisant la protéine Vpr couplée à la GFP (Green Fluorescent Protein) a permis de montrer que le CPI de HIV-1 co-localise avec les microtubules et est accumulé au niveau du centre dôorganisation des microtubules (MTOC) situé à la périphérie du noyau (McDonald et al., 2002).

Arrivé à proximité du noyau, le CPI doit franchir la membrane nucléaire pour atteindre l'ADN chromosomique.

La plupart des rétrovirus requièrent une division cellulaire pour se répliquer. L'entrée dans le noyau s'effectue alors pendant la mitose, lorsque la membrane nucléaire est rompue.

Seuls les rétrovirus du genre lentivirus tel que HIV-1 peuvent infecter les cellules ne se divisant pas ou les cellules quiescentes. Pour ces virus, le CPI entrerait dans le nucléoplasme par l'intermédiaire de pores nucléaires. La taille du CPI a un diamètre bien supérieur à la taille du canal central des pores nucléaires. L'entrée au noyau nécessite donc un transport actif. Des protéines virales et cellulaires peuvent contribuer au transport du CPI à travers la membrane nucléaire favorisant ainsi l'accès à la chromatine et assurant l'intégration dans des domaines particuliers de celle-ci. Des études ont permis d'identifier des signaux de localisation nucléaire (NLS) portés par 3 protéines virales différentes : IN, MA et Vpr. Ces NLS sont reconnues par des protéines cellulaires appelées importines et permettent un transport actif de molécules du cytoplasme vers le noyau. L'association de ces 3 séquences contribuerait à un ciblage nucléaire réussi (Bouyac-Bertoia et al., 2001; Sherman and Greene, 2002; Suzuki and Craigie, 2007). Il a été aussi montré que la présence d'une structure nommée « DNA Flap central » est un facteur contribuant à l'import nucléaire du CPI. Cette structure d'ADN à 3 brins, présente au centre du génome du HIV-1, est formée lors de la transcription inverse. Un ADN viral sans la séquence DNA flap se retrouve accumulé au voisinage de la membrane nucléaire (Zennou et al., 2000).

Le CPI pourrait aussi interagir avec un facteur nommé TNPO3 pour son transport dans le noyau. TNPO3 est une protéine de la famille des importines β et elle a été retrouvée en interaction avec l'IN d'une part, et avec le complexe CA-NC du virus HIV-1, d'autre part. TNPO3 est requise pour qu'il y ait infection par des virus du genre lentivirus. En effet, une déplétion de TNPO3 inhibe la réplication du virus HIV-1. Cette inhibition se fait après l'étape d'import nucléaire mais avant l'étape d'intégration (Cribier et al., 2011; Valle-Casuso et al., 2012).

Une fois le CPI entré dans le noyau, le cycle se poursuit par un processus appelé intégration virale.

## 1.2 L'intégration virale

### *1.2.1 Mécanisme*

Le processus d'intégration représente l'étape finale des phases précoces de l'infection. C'est une étape essentielle du cycle rétroviral qui aboutit à l'insertion stable de l'ADN viral dans le génome cellulaire. L'étape d'intégration est réalisée par l'enzyme virale Intégrase (IN). Elle se déroule en 4 phases :

1. Les séquences U3 et U5 des LTR possèdent, à leurs extrémités, des sites de reconnaissance appelés séquence « att » pour attachement. L'IN reconnait ces séquences et se fixe dessus.

2. Lô IN clive les extrémités 3ô grâce à son activité endonucléasique. Cette attaque nucléophile libère le dinucléotide 5ôP-GT (dans le cas du virus HIV-1) permettant dôobtenir le dinucléotide CAï OH libre à chaque extrémité 3ô.

Ces deux étapes se déroulent dans le cytoplasme de la cellule infectée. Le complexe est ensuite transloqué à lôintérieur du noyau.

3. Chaque groupement 3ô-OH généré suite à lôexcision, attaque un brin de lôADN cellulaire pour intégrer lôADN viral dans le génome cellulaire, via un mécanisme de trans-estérification. Côest lôétape de transfert de brins.

4. Les deux nucléotides non appariés en 5ô sont éliminés et les nucléotides manquants sont additionnés au niveau du site dôinsertion. Le processus dôintégration est achevé par des enzymes cellulaires de lôhôte permettant la réparation des brèches formées par lôinsertion de lôADN viral (Skalka and Katz, 2005). La réaction finale aboutit à la présence dôun ADN viral intégré appelé provirus, au sein de lôADN génomique avec duplication dôune courte séquence chromosomique de part et dôautre de la séquence virale intégrée. La taille de cette duplication est spécifique de chaque virus, 6 pb pour les virus ASLV et 5 pb pour HIV-1, 4 pour MLV.

Une telle intégration est dite concertée, désignant lôinsertion simultanée des deux extrémités virales sur un même site de lôADN génomique.

### *1.2.2 Sites dôintégration*

Le site dôinsertion de lôADN viral est crucial pour le virus. Lôintégration dans les régions transcriptionnellement actives peut favoriser lôexpression des gènes viraux facilitant ainsi la production de particules virales infectieuses. Lôintégration dans des régions transcriptionnellement réprimées est en défaveur de lôexpression des gènes viraux mais peut faciliter la latence virale. En outre, lôintégration rétrovirale peut avoir des conséquences importantes pour lôhôte car le provirus peut influencer lôexpression des gènes près desquels il est inséré. En effet, lôactivité des promoteurs viraux insérés peut modifier lôexpression des gènes cellulaires. Ce type dôinsertion peut conduire à des processus de tumorigenèse lorsque ces gènes sont des proto-oncogènes. Des études ont montré quôin vivo, les sites dôintégration rétrovirale ne sont pas totalement aléatoires, présentent une faible spécificité de séquence et que le choix du site dôintégration est spécifique du rétrovirus. Ainsi, HIV-1 (lentivirus) favorise une intégration dans les unités de transcription des gènes actifs sans préférence particulière pour les introns ou les exons. Les virus comme MLV ou PERV (gammarétrovirus) et les spumavirus sôintègrent préférentiellement autour des sites de début de transcription et dans les ilôts CpG (Moalic et al., 2006; Moalic et al., 2009). Les alpharétrovirus et les deltarétrovirus ont une faible préférence dôintégration dans les unités de transcription ou les ilôts CpG ; enfin, les bétarétrovirus ne montrent pas de préférence dans le

choix du site dôintégration (Desfarges and Ciuffi, 2010). Des protéines cellulaires, comme la protéine LEDGF/p75 (détaillé dans le paragraphe 3 p14), peuvent aussi orienter le choix du site dôintégration dans le cas du HIV-1(Ciuffi et al., 2005).

### *1.2.3 Les formes virales circulaires*

En lôabsence dôintégration, lôADN viral linéaire se circularise donnant naissance à des formes circulaires non réplicatives. Ces formes circulaires portent une ou deux LTRs juxtaposées et peuvent sôaccumuler dans le noyau (Delelis et al., 2008). Côest en dosant ce type de structures que lôon peut évaluer lôefficacité de lôintégration.

# 2 La protéine intégrase

Lôintégrase est une enzyme rétrovirale essentielle qui fixe les deux extrémités de lôADN viral linéaire et permet son insertion dans le chromosome de la cellule hôte.

## 2.1 Structure en domaines de lôintégrase

Lôintégrase est une protéine codée par lôextrémité du gène pol. Elle est produite à partir dôun précurseur Gag-Pol dont elle est séparée par action de la protéase virale. Elle comporte 3 domaines fonctionnels : un domaine N-terminal (NTD), un domaine central (CCD) et un domaine C-terminal (CTD). Lôintégrase de MLV est une protéine de 45 kDa composée de 408 acides aminés. Les intégrases du HIV-1 et de lôALV sont des protéines de 32 kDa composées de 288 et 286 acides aminés respectivement. Les intégrases du HIV-1 et les aviaires sont, de nos jours, les plus étudiées.

### *2.1.1 Le domaine N-terminal*

Le domaine N-terminal ou NTD contient 4 résidus conservés dans toutes les intégrases rétrovirales. Ces résidus sont représentés par 2 histidines et 2 cystéines. Ce motif appelé « HHCC », analogue à un doigt de zinc, fixe ce cation. Son rôle pourrait être de favoriser le processus de multimérisation de la protéine entière, processus clé pour lôactivité dôintégration. Ce domaine est nécessaire au processus de clivage des extrémités 3ô et de transfert de brins et joue un rôle dans la fixation de la protéine à lôADN viral (Jaskolski et al., 2009).

### *2.1.2 Le domaine central*

Le domaine central ou CCD est également nommé cî ur catalytique car il contient le site actif de lôenzyme. Ce site est constitué par une triade catalytique déterminée par le motif D,D(35)E. Le dernier résidu aspargine et le résidu glutamine sont séparés par 35 acides aminés. Ce motif

DD(35)E est hautement conservé parmi les intégrases rétrovirales et est indispensable pour lôactivité catalytique de la protéine. La présence de cations métalliques (Mg2+ et Mn2+) est strictement requise pour les activités des intégrases rétrovirales. Le domaine central est aussi impliqué dans lôinteraction avec les extrémités de lôADN viral et dans la dimérisation de la protéine (Jaskolski et al., 2009).

### *2.1.3 Le domaine C-terminal*

Le domaine C-terminal ou CTD constitue le domaine le moins conservé entre les intégrases rétrovirales. Ce domaine possède une forte capacité de liaison à lôADN mais cette fixation se fait de manière non spécifique (ADN cellulaire et ADN viral). Ce domaine intervient aussi dans la multimérisation de lôintégrase (Jaskolski et al., 2009).

## 2.2 Structure oligomérique de lôintégrase

Chaque domaine est capable de se structurer indépendamment du reste de la protéine. La présence des trois domaines et la multimérisation de la protéine sont essentielles pour que la réaction dôintégration ait lieu. La protéine serait active sous forme dimérique lors des étapes de clivage et sous forme tétramérique lors des étapes de transfert (Faure et al., 2005).

Jusquôà très récemment, seules des structures cristallographiques des domaines pris séparément ou des domaines pris deux à deux ont été résolues : les fragments comprenant le CCD + CTD pour le virus du sarcome de Rous (RSV) (Yang et al., 2000) et les fragments comprenant le CCD + CTD (Chen et al., 2000) ou le NTD + CCD (Wang et al., 2001) pour le HIV-1. Une étude récente, réalisée par cryo-microscopie, a permis de montrer un complexe stable entre un tetramère dôintégrase du HIV-1 associé à un cofacteur cellulaire : LEDGF/p75 (Michel et al., 2009). Dernièrement, la première structure cristallographique de lôintégrase entière du prototype humain des spumavirus (PFV) a été enfin résolue : elle montre lôIN sous forme tétramérique en complexe avec de lôADN viral (Hare et al., 2010; Maertens et al., 2010).

Une étude de notre laboratoire a mis en évidence une structure du domaine central de lôintégrase aviaire différente de celle qui était connue jusquôà présent (Ballandras et al., 2011). Cette nouvelle structure présente un sillon basique dans lequel pourrait se placer un acide nucléique simple brin. Les auteurs ont montré que le domaine central de la protéine pourrait se trouver sous deux formes structurales et ont émis lôhypothèse que chacune des structures interviendrait à des étapes différentes du cycle viral, ce qui reste à confirmer.

# 3 LEDGF/P75 : co-facteur cellulaire de lôintégrase du HIV-1

Plusieurs études ont montré que des protéines cellulaires ou co-facteurs jouent un rôle auxiliaire durant l'intégration rétrovirale. Certains ont la capacité dôinteragir avec lôADN viral comme les protéines HMGA1 (Hight Mobility Group protein A1), BAF (Barrier to Autointegration Factor) et dôautres se lient directement à lôIN du HIV-1 comme INI1 (INtegrase Interactor 1), EED (Embryonic Ectoderm Development), HSP60 (Heat shock protein 60), Rad18, et LEDGF/p75 (Lens Epithelial Derived Growth Factor/p75) (Turlure et al., 2004; Van Maele et al., 2006). Lôimplication de ces cofacteurs dans le processus intégratif nôest pas toujours bien connue mais parmi eux, LEDGF/p75 est lôun des plus étudié. Cette protéine joue un rôle fondamental dans la réplication des rétrovirus du genre lentivirus.

## 3.1 Présentation et organisation génomique

Contrairement à ce que son nom initial indique LEDGF/p75 (Lens Epithelial Derived Growth Factor/p75) nôest pas un facteur de croissance, pas plus quôil nôest spéciýque des cellules de lôépithélium du cristallin. La protéine LEDGF/p75, codée par le gène PSIP1, est une protéine appartenant à la famille des protéines Hepatoma derived growth factor Related Proteins (HRPs). Chez les mammifères, il existe 6 protéines HRPs : HDGF, HRP1, HRP2, HRP3, LEDGF/p52 et LEGDF/p75. Ces 6 protéines possèdent une forte homologie de séquence des 95 premiers résidus de la partie N-terminale appelée domaine PWWP. Seulement 2 de ces protéines, LEGDF/p75 et HRP2, contiennent un domaine particulier appelé IBD (IN Binding Domain) leur permettant dôinteragir avec des intégrases lentivirales. Cependant, à lôinverse de LEDGF/p75, HRP2 ne se fixe pas à la chromatine et ne permet pas de lier lôIN à la chromatine (Cherepanov et al., 2003; Cherepanov et al., 2004; Vanegas et al., 2005).

Lôépissage alternatif du précurseur ARNm de LEDGF/p75 permet lôexpression de deux isoformes, p75 et p52, qui diffèrent seulement par leur domaine C-terminal. LEDGF/p75 est une protéine de 530 acides aminés, exprimée à tous les stades de développement et dans une grande variété dôorganes et de tissus. Elle possède plusieurs domaines fonctionnels :

- un domaine N-terminal contenant le motif PWWP (pour Pro-Trp-Trp-Pro) très conservé, commun à dôautres protéines associées à la chromatine. Ce domaine est impliqué dans la croissance cellulaire et la transcription (Llano et al., 2006; Turlure et al., 2006)
- un signal de localisation nucléaire (NLS) permettant lôimport au noyau (Maertens et al., 2004; Vanegas et al., 2005).

- Deux copies d'un motif appelé crochet AT. Celui-ci se lie aux régions de l'ADN riches en Adénine et Thymine (Cherepanov et al., 2004).
- Un domaine de fixation à l'IN appelé IBD (Integrase-Binding Domain) d'environ 80 acides aminés, très conservé chez les vertébrés et qui se trouve dans la partie C-terminale de la protéine (Cherepanov et al., 2004; Vanegas et al., 2005).

## 3.2 Fonctions

**Liaison à la chromatine**

LEDGF/p75 est une protéine nucléaire, étroitement liée à la chromatine tout au long du cycle cellulaire (Cherepanov et al., 2003; Maertens et al., 2003; Llano et al., 2004a; Vanegas et al., 2005; Turlure et al., 2006). La liaison de LEDGF/p75 à la chromatine est rendue possible grâce à la coopération des domaines PWWP et des 2 copies du motif A/T (Turlure et al., 2006).

La liaison de LEDGF/p75 à l'ADN, *in vitro*, est possible grâce aux régions comprenant le NLS jusqu'au deux copies du motif A/T (Turlure et al., 2006).

**Régulation transcriptionnelle**

Il a été suggéré que LEDGF/p75 joue un rôle dans la régulation transcriptionnelle parce qu'elle a été trouvée comme interactant du co-activateur transcriptionnel PC4 isolé à partir d'extraits de cellules Hela (Ge et al., 1998).

LEDGF/p75 pourrait jouer un rôle dans l'activation des gènes de réponse au stress. Sa présence confère une protection contre l'apoptose de cellules sous stress en se fixant spécifiquement aux séquences HSE (éléments de réponse aux chocs thermiques) et STRE (éléments de réponse associés au stress) (Ganapathy et al., 2003; Shinohara et al., 2002).

## 3.3 LEDGF/p75 et Intégrase

LEDGF/p75 a été identifiée comme protéine associée à l'IN de HIV-1. Cherepanov et al. ont isolé, à partir des noyaux de cellules humaines, un complexe de 61 Å composé du tétramère de l'IN du HIV-1 auquel était associée une protéine cellulaire d'une masse moléculaire apparente de 75 kDa. Cette protéine a été identifiée par spectrométrie de masse et s'est avérée correspondre à LEDGF/p75 (Cherepanov et al., 2003). Cette interaction directe entre l'IN du HIV-1 et LEDGF/p75 a été confirmée (i) par précipitation du complexe *in vitro* par des méthodes de pull-down et de co-immunoprecipitation (Maertens et al., 2003; Cherepanov et al., 2003; Cherepanov et al., 2004; Busschots et al., 2005) et (ii) lors de la recherche d'interactants de l'IN du HIV-1 en double hybride (Cherepanov et al., 2003; Emiliani et al., 2005). Il a été aussi démontré que LEDGF/p75 est capable de se fixer spécifiquement sur les intégrases d'autres lentivirus tels que HIV-2, SIV, BIV, MVV, EIAV mais pas sur les intégrases des gammarétrovirus (MLV), des

alpharétrovirus (RSV) ou des deltarétrovirus (HTLV) (Busschots et al., 2005; Cherepanov, 2007; Llano et al., 2006; Emiliani et al., 2005).

Les domaines impliqués dans lôinteraction entre IN du HIV-1 et LEDGF/p75 sont :

✦ **Pour lôIN :**

Le CCD contient les déterminants nécessaires à lôinteraction avec la région IBD de LEDGF/p75 (Cherepanov et al., 2004; Turlure et al., 2006). La technique de mutagenèse dirigée a permis d'identifier les acides aminés de l'IN (W131, W132, I161 à E170) impliqués dans la liaison à LEDGF/p75, en particulier l'acide aminé Q168 (Emiliani et al., 2005; Busschots et al., 2007). La mutation Q168A n'affecte pas l'activité enzymatique de l'IN mais affecte son interaction avec LEDGF/p75.

Le NTD contribue à lôinteraction entre lôIN et LEDGF/p75. Des délétions ou des mutations dans ce domaine suppriment la distribution nucléaire et donc la colocalisation entre les 2 partenaires in vitro. De plus, ce domaine, analogue à un doigt de zinc, qui favorise le processus de multimérisation de lôIN, permettant dôaugmenter l'affinité de l'IN pour LEDGF/p75 (Maertens et al., 2003).

Le CTD de lôIN nôintervient pas dans lôinteraction avec LEDGF/p75.

✦ **Pour LEDGF** : Seul le domaine IBD de LEDGF/p75 intervient dans lôinteraction avec lôIN. Ce domaine est situé sur la partie C-terminale de lôisoforme p75 entre les résidus 340 et 417 (De Rijck et al., 2006; Cherepanov et al., 2004; Llano et al., 2006; Vanegas et al., 2005).

## 3.4 LEDGF/p75 dans le cycle viral

Un certain nombre dôétudes ont permis de montrer que LEDGF/p75 était nécessaire à la réplication du HIV-1.

**La transcription inverse :** Il a été montré que LEDGF/p75 nôa aucun effet sur lôétape de transcription inverse (Emiliani et al., 2005; Llano et al., 2006; De Rijck et al., 2006).

**Le complexe de préintégration** : LEDGF/p75 est un composant du CPI des virus HIV-1 et FIV. Le CPI peut être immunoprécipité à partir du cytoplasme de cellules avec des anticorps anti-LEDGF/p75 (Llano et al., 2004a). Cependant, LEDGF/p75 ne joue aucun rôle dans le transport du CPI jusquôà la périphérie du noyau (Emiliani et al., 2005; Llano et al., 2004a).

**Stabilisation et localisation cellulaire de lôintégrase :** LEDGF/p75 protège lôIN de la dégradation par le protéasome (Llano et al., 2004b; Emiliani et al., 2005; Llano et al., 2006). En

effet, en lôabsence de LEDGF/p75, lôIN des virus HIV-1 et FIV se révèlent être extrêmement instables et deviennent à peine détectables en cellules (Llano et al., 2004b).

En cellules, lôIN de HIV-1 se localise dans le noyau, associée à la chromatine durant le cycle cellulaire. Afin de comprendre le rôle de LEDGF/p75 dans la localisation cellulaire de lôIN *in cellulo*, la stratégie employée est lôabolition de lôinteraction entre lôIN et LEDGF/p75. Ceci est réalisé de plusieurs manières :

- en modifiant les domaines de lôIN ou de LEDGF/p75 impliqués dans cette liaison (par mutation du résidu Q168 sur lôIN ou par délétion du domaine IBD de LEDGF/p75) et observation de la localisation cellulaire des différents partenaires (Maertens et al., 2003; Emiliani et al., 2005; Busschots et al., 2007),
- en inhibant ou éteignant lôexpression de LEDGF/p75 par des approches ARN interférence ou knock-out respectivement, ou en délocalisant LEDGF/p75 par mutation de son NLS (Maertens et al., 2003; Maertens et al., 2004; Emiliani et al., 2005; Vanegas et al., 2005; Turlure et al., 2006; Shun et al., 2007).

De ces études, il a été observé quôen lôabsence de LEDGF/p75, lôexpression de lôIN est réduite et se retrouve localisée dans le cytoplasme des cellules et non plus essentiellement dans le noyau. De plus, celle-ci nôest plus associée aux chromosomes à lôinterphase ou lors de la mitose. A lôinverse, lors dôune surexpression de LEDGF/p75, on observe lôIN sur les chromosomes. LEDGF/p75 est donc capable dôinfluencer la rétention de lôIN dans le noyau des cellules et de cibler directement lôIN sur les chromosomes de la cellule hôte (Emiliani et al., 2005; Maertens et al., 2003).

**Les activités catalytiques de lôintégrase** : LEDGF/p75 augmente les activités enzymatiques (3ô processing, transfert de brins et intégration concertée) des INs des virus HIV-1, BIV, EIAV *in vitro* (Cherepanov, 2007). Il a été montré que, sous certaines conditions, LEDGF/p75 peut augmenter jusquôà 150 fois lôactivité de transfert de brins (Cherepanov et al., 2003; Cherepanov et al., 2004; Turlure et al., 2006).

**Intégration virale** : LEDGF/p75 joue un rôle au niveau de lôétape dôintégration. En effet, lôabsence de LEDGF/p75 diminue considérablement le nombre de copies intégrées (Emiliani et al., 2005; Ciuffi et al., 2005; Llano et al., 2006; De Rijck et al., 2006). Par ailleurs, lôanalyse des sites dôintégration du HIV-1 dans les cellules humaines déplétées en LEDGF/p75 a montré que lôintégration était : 1) moins fréquente dans les unités transcriptionnelles ; 2) moins fréquente dans les gènes régulés par LEDGF/p75 ; 3) plus fréquente dans les régions riches en GC (Ciuffi et al., 2005; Marshall et al., 2007). Le choix des sites dôintégration, en lôabsence de LEDGF/p75, ressemble alors à ceux observés avec lôIN du virus MLV.

A lôissu de toutes ces observations, un modèle du fonctionnement de LEDGF/p75 durant lôintégration par le HIV-1 a été proposé par Engelman et Cherepanov.

## 3.5 Cas de LEDGF/p52

La seconde isoforme p52 de 333 acides aminés, jouerait un rôle comme co-activateur général de la transcription avec une activité plus importante que p75 (Ge et al., 1998) mais nôinteragit pas avec lôIN dôHIV-1 *in vitro* et *in cellulo* (Ge et al., 1998; Maertens et al., 2003). Ne possédant pas de région IBD, LEDGF/p52 nôa aucune affinité pour lôIN du HIV-1 *in vitro* (Maertens et al., 2003). De plus, LEDGF/p52 ne protège pas lôIN de la dégradation par le protéasome (Llano et al., 2004b).

## 3.6 Conclusion

L'interaction de l'IN avec LEDGF/p75 est essentielle pour l'intégration et la réplication des virus du genre lentivirus. LEDGF/p75, composant du CPI, protège l'IN de lôaction du protéasome et augmente ses activités enzymatiques. LEDGF/p75, par son interaction avec lôIN, permet un ciblage nucléaire et chromosomique et oriente le choix des sites dôintégration dans les unités de transcription des gènes actifs. Ces résultats montrent combien LEDGF/p75 est un cofacteur cellulaire essentiel pour l'IN des lentivirus.

# Chapitre 3 : Les protéines à bromodomaines et Brd2

Une protéine, appelée Brd2 (pour bromodomain-containing protein 2), a été identifiée en interaction avec les IN des virus MLV et HIV-1 par la technique du double hybride en levure (Studamire and Goff, 2008). Mon travail ayant porté sur cette protéine, ce chapitre sera consacré à la présentation des protéines à bromodomaines et leurs interactions avec les virus. La protéine Brd2 étant retrouvé en interaction avec des protéines impliquées dans le remodelage de la chromatine, ce chapitre débutera par des généralités sur ce sujet.

## 1 Généralités et remodelage de la chromatine

Le noyau des cellules eucaryotes contient le support de l'information génétique qui est organisé en une structure complexe nommée chromatine, constituée d'ADN (35%) et de protéines (histones 35% et non histones 10 à 30%). La chromatine possède une structure hétérogène. C'est un enchevêtrement de fibres dont le diamètre varie non seulement au cours du cycle cellulaire mais aussi en fonction des régions chromosomiques observées. Il existe plusieurs niveaux de compaction de la chromatine avec des localisations nucléaires bien précises que l'on peut classer en deux grandes catégories (Fedorova and Zink, 2008):

- **l'hétérochromatine**, dense, est définie comme une structure qui ne change pas d'état de condensation au cours du cycle cellulaire. Elle est localisée principalement en périphérie du noyau et du nucléole. On distingue:

  - l'**hétérochromatine constitutive** qui contient peu de gènes, formée principalement de séquences répétées et dont les plus grandes régions sont situées à proximité des centromères et des télomères.

  - l'**hétérochromatine facultative** qui contient des régions codantes pouvant adopter les caractéristiques structurales et fonctionnelles de l'hétérochromatine, comme le chromosome X inactif chez la femelle des mammifères.

- **L'euchromatine**, claire, est décondensée. Elle est répartie à l'intérieur du nucléoplasme et contient les régions actives du génome.

Chaque noyau cellulaire contient environ deux mètres d'ADN ce qui nécessite une compaction de la molécule. Malgré ce degré de compaction, l'ADN subit des changements dynamiques au cours de nombreux processus génétiques, en particulier lors de sa réplication (cycle cellulaire), de l'expression des gènes, de la recombinaison et la réparation de l'ADN. Au

cours de ces processus, la chromatine doit être rapidement accessible aux protéines pour réguler ses fonctions.

Pour cela, la chromatine doit être organisée et structurée. Lôunité fondamentale de la chromatine est appelée le nucléosome : il est composé dôADN et dôhistones. Il représente le premier niveau de compaction de lôADN. Cette structure est ensuite régulièrement répétée pour former le nucléofilament souvent comparé à celle dôun collier de perles et qui peut, lui-même, adopter des niveaux d'organisation plus compacts.

Le nucléosome comprend une partie centrale et une région internucléosomique. La particule cî ur est composée de 146 paires de bases dôADN enroulées autour dôun octamère dôhistones constitué de deux dimères H2A et H2B et dôun tétramère H3 et H4.

Chaque histone possède, à ses extrémités NH2 et COOH terminales, des segments de 20 à 35 acides aminés, appelés queues qui sont particulièrement riches en résidus lysine. Dans le nucléosome, ces queues sont flexibles permettant ainsi lôaccès à des enzymes ou lôinteraction avec dôautres nucléosomes. Ces régions sont ainsi la cible de nombreuses modifications post-traductionnelles qui participent à la régulation de la compaction et de lôaccessibilité de la chromatine par des protéines régulatrices. Elles affectent leurs charges, lôaccessibilité à lôADN, les interactions protéines/protéines avec le nucléosome mais aussi lôexpression ou la répression des gènes. Il existe quatre types majeurs de modifications des histones : l'acétylation, la méthylation, la phosphorylation et lôubiquitinylation. Les différentes combinaisons de ces modifications forment un code appellé « code epigénétique des histones » (Sanchez and Zhou, 2009).

Parmi ces modifications, lôacétylation des lysines est le processus le plus dynamique car celle-ci dirige à la fois des changements structurels de la chromatine ainsi que la transcription des gènes. Cette modification est une marque de lôeuchromatine et sôopère sur les queues N-terminales des histones H3 et H4.

Lôacétylation consiste en l'addition d'un groupement acétyl CH3-C=O sur les résidus lysine en position N-terminale ou au sein de la chaîne polypeptidique des histones H3 ou H4. Cette réaction est catalysée par les Histones AcetylTransférases (HAT). Les lysines, chargées positivement, peuvent se coller à lôADN qui, lui, est chargé négativement. Lôacétylation des lysines neutralise cette charge positive et réduit lôaffinité entre lôADN et les histones intra- ou internucléosomes provoquant un relâchement du complexe. La chromatine, ainsi décondensée, est accessible à lôARN polymérase II et aux facteurs de transcription.

L'acétylation est une modification réversible et le clivage du groupement acétyl est catalysé par les histones déacétylases (HDAC). Cette modification conduit à la formation dôune chromatine condensée et est associée à la répression de la transcription.

# 2 Les protéines à bromodomaines

Le bromodomaine est un motif de séquence conservé, composé approximativement de 110 acides aminés. La caractéristique majeure du bromodomaine est sa reconnaissance spécifique et sa fixation aux lysines acétylées sur des queues N-terminales des histones (Florence & Faller 2001, Sanchez & Zhou 2009; Umehara, Nakamura, Jang, et al. 2010). Le génome humain code 42 protéines à bromodomaines, chacune contenant de 1 à 6 bromodomaines. Il existe 3 groupes majeurs de protéines contenant des bromodomaines :

(i) les histones acétyltransférases nucléaires (HATs) comme Gcn5, P/CAF, TAF1, CBP,

(ii) les facteurs de remodelage de la chromatine ATP-dépendant comme Swi2, Snf2,

(iii) les protéines de la famille des bromodomaines et extra-terminale domaine (BET).

# 3 Les protéines BET

Les protéines appartenant à la famille BET ont été trouvées chez les plantes, les champignons, les animaux et possèdent une architecture unique avec un (plantes) ou deux bromodomaines (BD) (animaux et levures) en tandem nommés BD1 et BD2 situés dans la partie N-terminale de la protéine. Les protéines BET possèdent toutes un NLS et un domaine extra-terminal (ET) conservé, trouvé dans la partie C-terminale et qui est impliqué dans des interactions protéine-protéine. Chez les mammifères, la famille des protéines BET comprend : Brd2/RING3 (bromodomain-containing protein 2), Brd3/ORFX, BRD6/BRDT (spécifiquement exprimé dans les testicules et dans les ovocytes) et Brd4. La protéine Brd4 comprend 2 variants d'épissage : un variant court Brd4/HUNK-1 et un variant long Brd4/MCAP.

Dans ce travail, nous nous intéressons plus particulièrement à la protéine Brd2 humaine.

# 4 La protéine Brd2 humaine

Le gène *brd2* a été initialement identifié dans la région du complexe majeur d'histocompatibilité de classe II situé sur le chromosome humain 6p21.3. Il est l'orthologue du gène fsh de la drosophile (Drosophila gene female sterile homeotic) et BDF1 chez la levure (S. cerevisiae gene bromodomain factor 1). Ce gène a été initialement nommé *ring3* (really interesting new gene-3) (Beck et al., 1992). Il a été par la suite rebaptisé *brd2*.

Bien que la protéine Brd2 se soit révélée être exprimée dans une grande variété de tissus, cette protéine est fortement exprimée dans les tissus avec des compartiments épithéliaux qui subissent un remodelage hormono-modulé comme la glande mammaire, l'utérus, l'épididyme, ainsi que les ovaires et les testicules (Rhee et al., 1998; Trousdale and Wolgemuth, 2004).

## 4.1 Fonctions de Brd2 :

*brd2* est un gène essentiel. Sa double mutation, chez la souris, conduit à une mortalité embryonnaire. Le phénotype observé suggère une fonction critique pour Brd2 dans le développement du tube neural et la différenciation (Shang et al. 2009). Il a été montré, toujours chez la souris, quôune perturbation du gène *brd2* dans sa région promotrice entrainant une expression réduite de Brd2 ne sôaccompagne pas dôanomalie de développement. Par contre, ces souris présentent une obésité importante accompagnée dôune hyperinsulinémie, mais avec une tolérance au glucose améliorée ainsi quôune faible glycémie (Wang et al., 2010)

Dans une étude réalisée par Greenwald et ses collaborateurs (Greenwald et al., 2004), il a été montré quôune surexpression de Brd2 spécifiquement dans des lymphocytes de souris provoque un développement de lymphomes à cellules B et des leucémies.

De plus, Brd2 joue un rôle dans le cycle cellulaire. Sur des cellules humaines 293T, une surexpression entraine une augmentation de la proportion de cellules en phase G1 et une diminution de la population en phase S du cycle cellulaire (Ottinger et al., 2006).

Brd2 a une activité sérine/thréonine kinase élevée dans les leucémies humaines. Brd2 présente les caractéristiques dôune kinase suite à des signaux mitotiques (Denis et al., 2000).

**Brd2 est impliquée dans la régulation de la transcription**.

Une étude menée par Denis et ses collaborateurs (Denis et al., 2006) a permis dôisoler un complexe nucléaire multiprotéique associé à Brd2 en utilisant une méthode de purification par affinité à partir de cellules cultivées. Les auteurs ont ensuite utilisé des méthodes protéomiques ainsi que de la spectrométrie de masse pour identifier les divers composants de ce complexe. Brd2 a été retrouvée en interaction avec plusieurs co-activateurs et co-répresseurs de la machinerie basale de transcription comme : TBP (TATA box binding protein), lôARN polymérase II, le facteur de transcription E2F, le complexe multiproteique Mediator, les protéines impliquées dans les modifications des histones (HAT et HDAC), les protéines impliquées dans le remodelage de la chromatine (SWI/SNF), et dôautres protéines X qui peuvent être impliquées dans cette interaction. Ce complexe multiprotéique est susceptible de contribuer à la régulation de la transcription à de multiples niveaux, à la régulation du cycle cellulaire et joue un rôle dans la prolifération et le cancer (Denis et al., 2006; Peng et al., 2007).

Brd2 interagit directement avec E2F-1 et avec TBP et forme ainsi un pont entre ces 2 facteurs de transcription. Peng J. et ses collaborateurs (Peng et al., 2007) ont montré que Brd2 sôassocie à TBP par lôintermédiaire de 26 acides aminés situés dans le premier bromodomaine de Brd2 et que Brd2 joue un rôle essentiel dans le recrutement de TBP au complexe de transcription E2F-1.

Il a été aussi montré que Brd2 transactive les promoteurs des gènes régulant le cycle cellulaire tels que la cycline D1, cycline A et cycline E (Denis et al., 2000; Greenwald et al., 2004; LeRoy et al., 2008; Ottinger et al., 2006).

## 4.2 Localisation cellulaire de Brd2.

Brd2 est principalement associée à l'euchromatine (Denis et al., 2000; Mattsson et al., 2002) Cette protéine reste associée à la chromatine lors de la compaction de l'ADN au cours de la mitose (Kanno et al., 2004) mais ceci n'est pas observé par tous les auteurs. Il a été montré que la protéine Brd2 était séquestrée dans le cytoplasme de cellules cultivées dans un milieu sans sérum mais transloquée au noyau lors d'une stimulation avec du sérum (Guo et al., 2000). Dans cette même étude, les auteurs ont démontré que le NLS de Brd2 était suffisant pour cibler la protéine au noyau des cellules.

**Liaison aux lysines acétylées.**

Dans un noyau intact, Brd2 est principalement associée à la lysine acétylée 12 de l'histone H4 (H4K12Ac) dans des lignées de cellules NIH/3T3 ou Hela (Kanno et al., 2004). Une autre étude indique que Brd2 est fortement liée à la chromatine contenant les lysines acétylées 12 ou 5 de l'histone H4 (H4K12Ac ou H4K5Ac). A l'inverse, Brd2 est à peine liée à la chromatine contenant la lysine acétylée 16 de l'histone H4 (H4K16Ac) ou la lysine acétylée 9 de l'histone H3 (H3K9Ac) (LeRoy et al., 2008). Récemment, les structures des bromodomaines BD1 et BD2 ont été résolues en association avec H4K5Ac/K12Ac (Huang et al., 2007; Umehara et al., 2010b; Umehara et al., 2010a)

# 5 Les protéines BET en interaction avec d'autres virus

Plusieurs études ont montré que les protéines à bromodomaines sont capables d'interagir avec des protéines virales ou d'interférer avec leur fonction.

## 5.1 Brd4 et Papillomavirus :

You et ses collaborateurs (You et al., 2004) ont montré que la partie C-terminale du variant long de Brd4/MCAP interagit avec la protéine E2 du papillomavirus bovin (BPV). La protéine E2 sert de ligand entre le génome épisomal du BPV et Brd4 qui est, lui-même, associé aux chromosomes mitotiques. Ceci permet au génome viral d'être transmis aux cellules filles.

## 5.2 Protéines BET et KSHV :

Brd2, Brd3 et les 2 variants de Brd4 interagissent avec LANA-1 (Latency-associated nuclear antigen 1), une protéine modulant la transcription et la réplication du génome épisomal du Kaposi's sarcoma-associated herpesvirus (KSHV). Ces interactions impliquent le domaine ET de ces protéines BET et la partie C-terminale de LANA-1 engagée dans la liaison aux chromosomes mitotiques de lôhôte (Viejo-Borbolla et al., 2005; You et al., 2006; Ottinger et al., 2006).

Ainsi, il a été montré que Brd2 colocalise avec LANA-1 dans des cellules transfectées avec un vecteur codant la protéine LANA-1 que ce soit dans des cellules en mitose ou à lôinterphase (Mattsson et al., 2002).

## 5.3 Protéines BET et autres herpes virus :

Comme le KSHV, le gamma herpes virus murin MHV-68 est aussi un gamma-2-herpes virus. Des souris infectées avec le MHV-68 servent de modèles pour lôinfection et la pathogénèse des gamma-herpes virus. Il a été montré que Brd2, Brd3 et Brd4/HUNK-1 interagissent avec la protéine MHV-68 orf 73, un homologue de LANA-1 (Ottinger et al., 2009). Les auteurs ont identifié les sites de liaison de la protéine MHV-68 orf 73 avec les BRDs et ils ont montré quôune mutation dans ces sites diminue lôassociation de la protéine MHV-68 orf 73 avec la chromatine et lôempêche dôactiver les promoteurs de la cycline D1, D2 et E.

Le virus dôEpstein-Barr (EBV) est un gamma-herpes virus lointainement lié au KSHV. La protéine EBNA1 (Epstein-Barr virus Nuclear Antigen 1) est un homologue fonctionnel de LANA et joue un rôle dans la persistance, la réplication et la régulation de la transcription. Récemment, il a été montré que Brd4 interagit avec EBNA1 et que cette interaction est importante pour lôactivation de la transcription dôéléments du génome viral (Lin et al., 2008).

## 5.4 Protéines BET et rétrovirus :

Cho et ses collaborateurs ont montré que la protéine Tax du virus T-lymphotropique humain (HTLV-1) et Brd4 entrent en compétition pour le recrutement dôun facteur positif dôélongation de la transcription (P-TEFb). En effet, lôinteraction entre Tax et P-TEFb est essentielle pour la transactivation des promoteurs LTR (Cho et al., 2007).

De la même manière que Tax, il existe une compétition entre la protéine Tat du HIV et Brd4 pour le recrutement du facteur P-TEFb (Yang et al., 2005).

## 5.5 Conclusion

Toutes ces observations montrent combien les protéines virales peuvent exploiter le fait que les protéines BET participent à la régulation de l'expression des gènes cellulaires. Certains virus semblent même usurper ces protéines BET pour des fonctions assez inhabituelles comme le fait d'attacher le génome épisomal aux chromosomes mitotiques de l'hôte (pour E2 du BPV et LANA-1 du KSHV)

## 5.6 Brd2 et Intégrase

Une étude récente a montré que la protéine Brd2 est un partenaire potentiel de l'intégrase (IN). Ce résultat a été obtenu par un criblage double hybride et en GST pull down, réalisé sur l'IN du Moloney Murine Leukemia Virus (MoMLV) contre une banque d'ADNc murins (Studamire and Goff, 2008). Cette étude a permis d'isoler 9 séquences de *brd2* murin (2 isolats appartenant au domaine BD2 et 7 isolats au domaine C-terminal de Brd2) interagissant avec l'IN du MoMLV et plus précisément avec la partie comprenant le CCD et CTD de l'IN. Vingt-sept autres protéines ont été identifiées et validées comme partenaires cellulaires de l'IN dont certaines, comme Brd2, ont une activité de régulation de la transcription. Dans cette même étude, les auteurs ont aussi montré, par la technique du double hybride, que l'IN du HIV-1 interagit également avec Brd2 mais bien plus faiblement qu'avec l'IN MoMLV. En revanche, *in vitro*, l'intensité de l'interaction est égale pour les deux intégrases. Ces résultats permettent de suggérer que Brd2 pourrait jouer un rôle dans l'intégration de ces deux rétrovirus.

# Références

**Ballandras, A., Moreau, K., Robert, X., Confort, M.-P., Merceron, R., Haser, R., Ronfort, C. and Gouet, P.** (2011). A crystal structure of the catalytic core domain of an avian sarcoma and leukemia virus integrase suggests an alternate dimeric assembly. *PLoS ONE* **6**, e23032.

**Beck, S., Hanson, I., Kelly, A., Pappin, D. J. and Trowsdale, J.** (1992). A homologue of the Drosophila female sterile homeotic (fsh) gene in the class II region of the human MHC. *DNA Seq.* **2**, 203ï 210.

**Bouyac-Bertoia, M., Dvorin, J. D., Fouchier, R. A., Jenkins, Y., Meyer, B. E., Wu, L. I., Emerman, M. and Malim, M. H.** (2001). HIV-1 infection requires a functional integrase NLS. *Mol. Cell* **7**, 1025ï 1035.

**Busschots, K., Vercammen, J., Emiliani, S., Benarous, R., Engelborghs, Y., Christ, F. and Debyser, Z.** (2005). The interaction of LEDGF/p75 with integrase is lentivirus-specific and promotes DNA binding. *J. Biol. Chem.* **280**, 17841ï 17847.

**Busschots, K., Voet, A., De Maeyer, M., Rain, J.-C., Emiliani, S., Benarous, R., Desender, L., Debyser, Z. and Christ, F.** (2007). Identification of the LEDGF/p75 binding site in HIV-1 integrase. *J. Mol. Biol.* **365**, 1480ï 1492.

**Chen, J. C., Krucinski, J., Miercke, L. J., Finer-Moore, J. S., Tang, A. H., Leavitt, A. D. and Stroud, R. M.** (2000). Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8233ï 8238.

**Cherepanov, P.** (2007). LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro. *Nucleic Acids Res.* **35**, 113ï 124.

**Cherepanov, P., Maertens, G., Proost, P., Devreese, B., Van Beeumen, J., Engelborghs, Y., De Clercq, E. and Debyser, Z.** (2003). HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J. Biol. Chem* **278**, 372ï 381.

**Cherepanov, P., Devroe, E., Silver, P. A. and Engelman, A.** (2004). Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. *J. Biol. Chem.* **279**, 48883ï 48892.

**Cho, W.-K., Zhou, M., Jang, M. K., Huang, K., Jeong, S.-J., Ozato, K. and Brady, J. N.** (2007). Modulation of the Brd4/P-TEFb interaction by the human T-lymphotropic virus type 1 tax protein. *J. Virol.* **81**, 11179ï 11186.

**Ciuffi, A., Llano, M., Poeschla, E., Hoffmann, C., Leipzig, J., Shinn, P., Ecker, J. R. and Bushman, F.** (2005). A role for LEDGF/p75 in targeting HIV DNA integration. *Nat. Med.* **11**, 1287ï 1289.

**Cribier, A., Ségéral, E., Delelis, O., Parissi, V., Simon, A., Ruff, M., Benarous, R. and Emiliani, S.** (2011). Mutations affecting interaction of integrase with TNPO3 do not prevent HIV-1 cDNA nuclear import. *Retrovirology* **8**, 104.

**De Rijck, J., Vandekerckhove, L., Gijsbers, R., Hombrouck, A., Hendrix, J., Vercammen, J., Engelborghs, Y., Christ, F. and Debyser, Z.** (2006). Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. *J. Virol* **80**, 11498ï 11509.

**Delelis, O., Carayon, K., Saïb, A., Deprez, E. and Mouscadet, J.-F.** (2008). Integrase and integration: biochemical activities of HIV-1 integrase. *Retrovirology* **5**, 114.

**Denis, G. V., Vaziri, C., Guo, N. and Faller, D. V.** (2000). RING3 kinase transactivates promoters of cell cycle regulatory genes through E2F. *Cell Growth Differ.* **11**, 417ï 424.

**Denis, G. V., McComb, M. E., Faller, D. V., Sinha, A., Romesser, P. B. and Costello, C. E.** (2006). Identification of transcription complexes that contain the double bromodomain protein Brd2 and chromatin remodeling machines. *J. Proteome Res* **5**, 502ï 511.

**Denner, J.** (2008). Recombinant porcine endogenous retroviruses (PERV-A/C): a new risk for xenotransplantation? *Arch. Virol.* **153**, 1421ï 1426.

**Denner, J.** (2011). Infectious risk in xenotransplantation--what post-transplant screening for the human recipient? *Xenotransplantation* **18**, 151ï 157.

**Desfarges, S. and Ciuffi, A.** (2010). Retroviral integration site selection. *Viruses* **2**, 111ï 130.

**Emiliani, S., Mousnier, A., Busschots, K., Maroun, M., Van Maele, B., Tempé, D., Vandekerckhove, L., Moisant, F., Ben-Slama, L., Witvrouw, M., et al.** (2005). Integrase mutants defective for interaction with LEDGF/p75 are impaired in chromosome tethering and HIV-1 replication. *J. Biol. Chem* **280**, 25517ï 25523.

**Farnet, C. M. and Bushman, F. D.** (1997). HIV-1 cDNA integration: requirement of HMG I(Y) protein for function of preintegration complexes in vitro. *Cell* **88**, 483ï 492.

**Faure, A., Calmels, C., Desjobert, C., Castroviejo, M., Caumont-Sarcos, A., Tarrago-Litvak, L., Litvak, S. and Parissi, V.** (2005). HIV-1 integrase crosslinked oligomers are active in vitro. *Nucleic Acids Res.* **33**, 977ï 986.

**Fedorova, E. and Zink, D.** (2008). Nuclear architecture and gene regulation. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1783**, 2174ï 2184.

**Florence, B. and Faller, D. V.** (2001). You bet-cha: a novel family of transcriptional regulators. *Front. Biosci.* **6**, D1008ï 1018.

**Freed, E. O.** (2001). HIV-1 replication. *Somat. Cell Mol. Genet.* **26**, 13ï 33.

**Ganapathy, V., Daniels, T. and Casiano, C. A.** (2003). LEDGF/p75: a novel nuclear autoantigen at the crossroads of cell survival and apoptosis. *Autoimmunity Reviews* **2**, 290ï 297.

**Ge, H., Si, Y. and Roeder, R. G.** (1998). Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. *EMBO J.* **17**, 6723ï 6729.

**Greenwald, R. J., Tumang, J. R., Sinha, A., Currier, N., Cardiff, R. D., Rothstein, T. L., Faller, D. V. and Denis, G. V.** (2004). Eμ-BRD2 transgenic mice develop B-cell lymphoma and leukemia. *Blood* **103**, 1475 ï 1484.

**Guo, N., Faller, D. V. and Denis, G. V.** (2000). Activation-induced nuclear translocation of RING3. *J. Cell. Sci.* **113 ( Pt 17)**, 3085ï 3091.

**Hare, S., Gupta, S. S., Valkov, E., Engelman, A. and Cherepanov, P.** (2010). Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* **464**, 232ï 236.

**Huang, H., Zhang, J., Shen, W., Wang, X., Wu, J., Wu, J. and Shi, Y.** (2007). Solution structure of the second bromodomain of Brd2 and its specific interaction with acetylated histone tails. *BMC Struct. Biol.* **7**, 57.

**Jaskolski, M., Alexandratos, J. N., Bujacz, G. and Wlodawer, A.** (2009). Piecing together the structure of retroviral integrase, an important target in AIDS therapy. *FEBS J.* **276**, 2926ï 2946.

**Kanno, T., Kanno, Y., Siegel, R. M., Jang, M. K., Lenardo, M. J. and Ozato, K.** (2004). Selective recognition of acetylated histones by bromodomain proteins visualized in living cells. *Mol. Cell* **13**, 33ï 43.

**Kukolj, G., Katz, R. A. and Skalka, A. M.** (1998). Characterization of the nuclear localization signal in the avian sarcoma virus integrase. *Gene* **223**, 157ï 163.

**LeRoy, G., Rickards, B. and Flint, S. J.** (2008). The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. *Mol. Cell* **30**, 51ï 60.

**Li, L., Li, H. S., Pauza, C. D., Bukrinsky, M. and Zhao, R. Y.** (2005). Roles of HIV-1 auxiliary proteins in viral pathogenesis and host-pathogen interactions. *Cell Res.* **15**, 923ï 934.

**Lin, C.-W. and Engelman, A.** (2003). The barrier-to-autointegration factor is a component of functional human immunodeficiency virus type 1 preintegration complexes. *J. Virol.* **77**, 5030ï 5036.

**Lin, A., Wang, S., Nguyen, T., Shire, K. and Frappier, L.** (2008). The EBNA1 protein of Epstein-Barr virus functionally interacts with Brd4. *J. Virol.* **82**, 12009ï 12019.

**Llano, M., Vanegas, M., Fregoso, O., Saenz, D., Chung, S., Peretz, M. and Poeschla, E. M.** (2004a). LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes. *J. Virol.* **78**, 9524ï 9537.

**Llano, M., Delgado, S., Vanegas, M. and Poeschla, E. M.** (2004b). Lens epithelium-derived growth factor/p75 prevents proteasomal degradation of HIV-1 integrase. *J. Biol. Chem.* **279**, 55570ï 55577.

**Llano, M., Saenz, D. T., Meehan, A., Wongthida, P., Peretz, M., Walker, W. H., Teo, W. and Poeschla, E. M.** (2006). An essential role for LEDGF/p75 in HIV integration. *Science* **314**, 461ï 464.

**Maertens, G., Cherepanov, P., Pluymers, W., Busschots, K., De Clercq, E., Debyser, Z. and Engelborghs, Y.** (2003). LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *J. Biol. Chem* **278**, 33528ï 33539.

**Maertens, G., Cherepanov, P., Debyser, Z., Engelborghs, Y. and Engelman, A.** (2004). Identification and characterization of a functional nuclear localization signal in the HIV-1 integrase interactor LEDGF/p75. *J. Biol. Chem.* **279**, 33421ï 33429.

**Maertens, G. N., Hare, S. and Cherepanov, P.** (2010). The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature* **468**, 326ï 329.

**Marshall, H. M., Ronen, K., Berry, C., Llano, M., Sutherland, H., Saenz, D., Bickmore, W., Poeschla, E. and Bushman, F. D.** (2007). Role of PSIP1/LEDGF/p75 in Lentiviral Infectivity and Integration Targeting. *PLoS ONE* **2**,.

**Mattsson, K., Kiss, C., Platt, G. M., Simpson, G. R., Kashuba, E., Klein, G., Schulz, T. F. and Szekely, L.** (2002). Latent nuclear antigen of Kaposiôs sarcoma herpesvirus/human herpesvirus-8 induces and relocates RING3 to nuclear heterochromatin regions. *J. Gen. Virol.* **83**, 179ï 188.

**McDonald, D., Vodicka, M. A., Lucero, G., Svitkina, T. M., Borisy, G. G., Emerman, M. and Hope, T. J.** (2002). Visualization of the intracellular behavior of HIV in living cells. *J. Cell Biol.* **159**, 441ï 452.

**Michel, F., Crucifix, C., Granger, F., Eiler, S., Mouscadet, J.-F., Korolev, S., Agapkina, J., Ziganshin, R., Gottikh, M., Nazabal, A., et al.** (2009). Structural basis for HIV-1 DNA integration in the human genome, role of the LEDGF/P75 cofactor. *EMBO J.* **28**, 980ï 991.

**Miller, M. D., Farnet, C. M. and Bushman, F. D.** (1997). Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *Journal of Virology* **71**, 5382.

**Moalic, Y., Blanchard, Y., Félix, H. and Jestin, A.** (2006). Porcine endogenous retrovirus integration sites in the human genome: features in common with those of murine leukemia virus. *J. Virol.* **80**, 10980ï 10988.

**Moalic, Y., Félix, H., Takeuchi, Y., Jestin, A. and Blanchard, Y.** (2009). Genome areas with high gene density and CpG island neighborhood strongly attract porcine endogenous retrovirus for integration and favor the formation of hot spots. *J. Virol.* **83**, 1920ï 1929.

**Moreau, K., Faure, C., Verdier, G. and Ronfort, C.** (2002). Analysis of conserved and non-conserved amino acids critical for ALSV (Avian leukemia and sarcoma viruses) integrase functions in vitro. *Arch. Virol.* **147**, 1761ï 1778.

**Moreau, K., Faure, C., Violot, S., Verdier, G. and Ronfort, C.** (2003). Mutations in the C-terminal domain of ALSV (Avian Leukemia and Sarcoma Viruses) integrase alter the concerted DNA integration process in vitro. *Eur. J. Biochem.* **270**, 4426ï 4438.

**Moreau, K., Faure, C., Violot, S., Gouet, P., Verdier, G. and Ronfort, C.** (2004). Mutational analyses of the core domain of Avian Leukemia and Sarcoma Viruses integrase: critical residues for concerted integration and multimerization. *Virology* **318**, 566ï 581.

**Mornex, J. F.** (2000). [Endogenous porcine retroviruses and xenotransplantation]. *Pathol. Biol.* **48**, 395ï 398.

**Nakamura, Y., Umehara, T., Nakano, K., Jang, M. K., Shirouzu, M., Morita, S., Uda-Tochio, H., Hamana, H., Terada, T., Adachi, N., et al.** (2007). Crystal structure of the human BRD2 bromodomain: insights into dimerization and recognition of acetylated histone H4. *J. Biol. Chem.* **282**, 4193ï 4201.

**Ottinger, M., Christalla, T., Nathan, K., Brinkmann, M. M., Viejo-Borbolla, A. and Schulz, T. F.** (2006). Kaposiôs Sarcoma-Associated Herpesvirus LANA-1 Interacts with the Short Variant of BRD4 and Releases Cells from a BRD4- and BRD2/RING3-Induced G1 Cell Cycle Arrest. *Journal of Virology* **80**, 10772ï10786.

**Ottinger, M., Pliquet, D., Christalla, T., Frank, R., Stewart, J. P. and Schulz, T. F.** (2009). The interaction of the gammaherpesvirus 68 orf73 protein with cellular BET proteins affects the activation of cell cycle promoters. *J. Virol.* **83**, 4423ï4434.

**Peng, J., Dong, W., Chen, L., Zou, T., Qi, Y. and Liu, Y.** (2007). Brd2 is a TBP-associated protein and recruits TBP into E2F-1 transcriptional complex in response to serum stimulation. *Mol. Cell. Biochem.* **294**, 45ï54.

**Platt, G. M., Simpson, G. R., Mittnacht, S. and Schulz, T. F.** (1999). Latent nuclear antigen of Kaposiôs sarcoma-associated herpesvirus interacts with RING3, a homolog of the Drosophila female sterile homeotic (fsh) gene. *J. Virol.* **73**, 9789ï9795.

**Rhee, K., Brunori, M., Besset, V., Trousdale, R. and Wolgemuth, D. J.** (1998). Expression and potential role of Fsrg1, a murine bromodomain-containing homologue of the Drosophila gene female sterile homeotic. *J. Cell. Sci.* **111 ( Pt 23)**, 3541ï3550.

**Sanchez, R. and Zhou, M.-M.** (2009). The role of human bromodomains in chromatin biology and gene transcription. *Curr Opin Drug Discov Devel* **12**, 659ï665.

**Shang, E., Wang, X., Wen, D., Greenberg, D. A. and Wolgemuth, D. J.** (2009). Double bromodomain-containing gene Brd2 is essential for embryonic development in mouse. *Dev. Dyn.* **238**, 908ï917.

**Sherman, M. P. and Greene, W. C.** (2002). Slipping through the door: HIV entry into the nucleus. *Microbes Infect.* **4**, 67ï73.

**Shinohara, T., Singh, D. P. and Fatma, N.** (2002). LEDGF, a survival factor, activates stress-related genes. *Prog Retin Eye Res* **21**, 341ï358.

**Shun, M.-C., Raghavendra, N. K., Vandegraaff, N., Daigle, J. E., Hughes, S., Kellam, P., Cherepanov, P. and Engelman, A.** (2007). LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev.* **21**, 1767ï1778.

**Skalka, A. M. and Katz, R. A.** (2005). Retroviral DNA integration and the DNA damage response. *Cell Death Differ.* **12 Suppl 1**, 971ï978.

**Studamire, B. and Goff, S. P.** (2008). Host proteins interacting with the Moloney murine leukemia virus integrase: multiple transcriptional regulators and chromatin binding factors. *Retrovirology* **5**, 48.

**Suzuki, Y. and Craigie, R.** (2007). The road to chromatin - nuclear entry of retroviruses. *Nat. Rev. Microbiol.* **5**, 187ï196.

**Trousdale, R. K. and Wolgemuth, D. J.** (2004). Bromodomain containing 2 (Brd2) is expressed in distinct patterns during ovarian folliculogenesis independent of FSH or GDF9 action. *Mol. Reprod. Dev.* **68**, 261ï268.

**Turlure, F., Devroe, E., Silver, P. A. and Engelman, A.** (2004). Human cell proteins and human immunodeficiency virus DNA integration. *Front. Biosci.* **9**, 3187ï 3208.

**Turlure, F., Maertens, G., Rahman, S., Cherepanov, P. and Engelman, A.** (2006). A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo. *Nucleic Acids Res* **34**, 1653ï 1665.

**Umehara, T., Nakamura, Y., Jang, M. K., Nakano, K., Tanaka, A., Ozato, K., Padmanabhan, B. and Yokoyama, S.** (2010a). Structural basis for acetylated histone H4 recognition by the human BRD2 bromodomain. *J. Biol. Chem* **285**, 7610ï 7618.

**Umehara, T., Nakamura, Y., Wakamori, M., Ozato, K., Yokoyama, S. and Padmanabhan, B.** (2010b). Structural implications for K5/K12-di-acetylated histone H4 recognition by the second bromodomain of BRD2. *FEBS Lett* **584**, 3901ï 3908.

**Valle-Casuso, J. C., Di Nunzio, F., Yang, Y., Reszka, N., Lienlaf, M., Arhel, N., Perez, P., Brass, A. L. and Diaz-Griffero, F.** (2012). TNPO3 Is Required for HIV-1 Replication after Nuclear Import but prior to Integration and Binds the HIV-1 Core. *J. Virol.* **86**, 5931ï 5936.

**Van Maele, B., Busschots, K., Vandekerckhove, L., Christ, F. and Debyser, Z.** (2006). Cellular co-factors of HIV-1 integration. *Trends Biochem. Sci.* **31**, 98ï 105.

**Vanegas, M., Llano, M., Delgado, S., Thompson, D., Peretz, M. and Poeschla, E.** (2005). Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering. *J. Cell. Sci.* **118**, 1733ï 1743.

**Viejo-Borbolla, A., Ottinger, M., Brüning, E., Bürger, A., König, R., Kati, E., Sheldon, J. A. and Schulz, T. F.** (2005). Brd2/RING3 interacts with a chromatin-binding domain in the Kaposiôs Sarcoma-associated herpesvirus latency-associated nuclear antigen 1 (LANA-1) that is required for multiple functions of LANA-1. *J. Virol.* **79**, 13618ï 13629.

**Wang, J. Y., Ling, H., Yang, W. and Craigie, R.** (2001). Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. *EMBO J.* **20**, 7333ï 7343.

**Wang, F., Liu, H., Blanton, W. P., Belkina, A., Lebrasseur, N. K. and Denis, G. V.** (2010). Brd2 disruption in mice causes severe obesity without Type 2 diabetes. *Biochem. J.* **425**, 71ï 83.

**Weiss, R. A.** (2006). The discovery of endogenous retroviruses. *Retrovirology* **3**, 67.

**Yang, Z. N., Mueser, T. C., Bushman, F. D. and Hyde, C. C.** (2000). Crystal structure of an active two-domain derivative of Rous sarcoma virus integrase. *J. Mol. Biol.* **296**, 535ï 548.

**Yang, Z., Yik, J. H. N., Chen, R., He, N., Jang, M. K., Ozato, K. and Zhou, Q.** (2005). Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol. Cell* **19**, 535ï 545.

**You, J., Croyle, J. L., Nishimura, A., Ozato, K. and Howley, P. M.** (2004). Interaction of the Bovine Papillomavirus E2 Protein with Brd4 Tethers the Viral DNA to Host Mitotic Chromosomes. *Cell* **117**, 349ï 360.

**You, J., Srinivasan, V., Denis, G. V., Harrington, W. J., Jr, Ballestas, M. E., Kaye, K. M. and Howley, P. M.** (2006). Kaposiôs sarcoma-associated herpesvirus latency-associated nuclear antigen interacts with bromodomain protein Brd4 on host mitotic chromosomes. *J. Virol.* **80**, 8909ï 8919.

**Zennou, V., Petit, C., Guetard, D., Nerhbass, U., Montagnier, L. and Charneau, P.** (2000). HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell* **101**, 173ï 185.