



**HAL**  
open science

# Étude comparative des processus intégratifs des rétrovirus aviaires et porcins

Elsy Al Andary

► **To cite this version:**

Elsy Al Andary. Étude comparative des processus intégratifs des rétrovirus aviaires et porcins. Sciences agricoles. Université Claude Bernard - Lyon I, 2011. Français. NNT : 2011LYO10347 . tel-00866109

**HAL Id: tel-00866109**

**<https://theses.hal.science/tel-00866109>**

Submitted on 26 Sep 2013

**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.

THESE DE L'UNIVERSITE DE LYON

Délivrée par

L'UNIVERSITE CLAUDE BERNARD LYON1

Ecole doctorale Evolution Ecosystèmes Microbiologie Modélisation

# ***DIPLOME DE DOCTORAT***

(arrêté du 7 août 2006)

*Soutenu publiquement le 19 décembre 2011*

*par*

**AL ANDARY Elsy**

**Etude comparative des processus intégratifs  
des rétrovirus aviaires et porcins**

## **Jury de soutenance**

Dr Y. CHEBLOUNE	Rapporteur
Dr M. LAVIGNE	Rapporteur
Dr F. LEFÈVRE	Rapporteur
Pr P. GOUET	Membre
Dr RONFORT Corinne	Directrice de thèse
Dr BLANCHARD Yannick	Co-Directeur de thèse

## *Remerciements*

J'adresse mes remerciements au Pr JF. Mornex et A. Jestin pour m'avoir accueillie au sein de leurs laboratoires. Je remercie également Dr C. Ronfort et Dr Y. Blanchard pour leur accueil au sein de leurs équipes.

Mes sincères remerciements vont à tous les membres de l'UGVB pour leur précieuse aide et leurs amitiés. Un merci va également à l'UMR754.

Je souhaite également remercier Dr A. Chaboud, Y. Tauran et Dr K. Moreau pour leurs conseils scientifiques.

Je remercie les membres du jury pour leur participation et le temps qu'ils ont consacré à cette thèse.

*A ma famille et mes amis,*

Tous mes remerciements à Dalida El Khoury et à ma meilleure amie Mira pour votre chaleureux accueil toutes les fois que j'en avais besoin.

Abdou Akkouche, Hanan Saad, Zeinab Hamzé, Nacira et Amal Abbassi... Je vous remercie pour le bon temps qu'on a passé ensemble durant ces années de thèse.

Antonin Demange, Salman Zeitouni, et la famille Chemali... sans vous ca aurait été dur en Bretagne. Merci pour votre aide et soutien.

Merci à Joe Obk et Michel Serhal pour leur aide précieuse. Merci également à Mirna et Ghina pour votre soutien et votre encouragement. Votre amitié m'a aidée à tenir bon.

A toi Virginie Gay, et à tes parents ... un GRAND merci, pour toutes les fois que vous m'aviez accueillie, écoutée et soutenue... J'en suis infiniment reconnaissante.

A mes parents, Nada et Nazih... A mes frères, Charbel et Rabih ... Je ne trouverai jamais assez de mots pour vous remercier! ... Heureusement, vous savez ce qu'il y a au fond de mon cœur.

*Je dédie cette thèse à mon père.*

## Résumé

Les rétrovirus sont des virus à ARN, enveloppés présents dans de nombreuses espèces animales de rente, chez les animaux de compagnie et chez l'homme. Une des particularités des rétrovirus concerne l'intégration du génome viral au sein du génome de la cellule infectée; cette intégration est réalisée par une enzyme virale, l'intégrase. Le projet de cette thèse vise à mieux comprendre le fonctionnement de cette enzyme notamment en identifiant des facteurs cellulaires interagissant avec celle-ci, facteurs qui pourraient être des agents favorisant le processus intégratif ou, au contraire, des agents restrictifs.

Les intégrases de deux modèles de rétrovirus ont été utilisées dans cette étude : L'intégrase de RAV1, un rétrovirus exogène aviaire du genre des alpharétrovirus appartenant au sous-groupe A de la famille des ASLV. Cette enzyme virale est largement étudiée soit au niveau structural ou fonctionnel, mais les données concernant ses partenaires cellulaires sont rares et insuffisantes. La seconde intégrase est celle du PERV A/C, un rétrovirus endogène porcine du genre gammarétrovirus. Aucune information sur cette enzyme n'a été décrite jusqu'à présent.

Ces deux enzymes, en fusion avec une étiquette 6xHistidine, ont été donc produites en bactérie, et en cellules d'insecte puis purifiées sur colonne d'affinité en FPLC. Leurs activités catalytiques ont été testées *in vitro*. Ces tests permettent de valoriser la capacité de l'intégrase à exercer principalement les 2 fonctions dont elle est responsable *in vivo*, le clivage en 3' et le transfert de brins, et une activité qu'elle exerce exclusivement *in vitro*, la désintégration. Les protéines pures et actives ont ensuite servies à la vérification de leur interaction avec une protéine cellulaire, Brd2. La technique 'Far western blot' a ainsi permis de valider l'interaction entre l'intégrase de PERV et la protéine cellulaire, puis d'identifier les domaines de l'intégrase et de Brd2 impliqués dans cette interaction.

A terme, l'identification de ce facteur cellulaire et la validation de son rôle dans le processus intégratif permettront de mieux comprendre ce processus particulier développé par les rétrovirus et pourront conduire au développement d'inhibiteurs dirigés contre cette interaction.

## Abstract

A critical step for retroviral replication is the stable integration of the provirus genome into the genome of its host; this integration is realized by a viral enzyme, the integrase. The aim of this work was to better understand the functioning of the integrase, particularly, by identifying host factors that might interact with it, and which could be factors favoring the integration process or, restrictive factors.

Therefore, we used two models of retroviral integrases: The integrase of RAV1, an alpharetrovirus belonging to the subgroup A of the family of ALSV. Although this viral enzyme is widely studied, still not enough data are available about its cellular cofactors. The second enzyme studied here is the integrase of PERV, a gammaretrovirus. No studies of either PERV integrase activities *in vitro* or of proteins interacting with this viral enzyme have been available until now.

In the present study, we have expressed the PERV and ALSV integrases as fusion proteins with a 6xHistidine Tag in both *Escherichia coli* and insect SF9 cells. After that, we analysed their ability to mediate catalytic activities (3'-end processing, strand transfer and disintegration) *in vitro*. We also investigated the interaction of these two viral enzymes with the cellular protein Brd2, using the Far western blot method. Our results validate Brd2 as a cofactor of PERV integrase and point to the important role of particular domains of the PERV integrase and Brd2 in mediating the interaction.

Finally, this study contribute to a better understanding of the precise interaction between cellular proteins and integrase, and may lead in the future to the development of protein-protein interaction inhibitors.

<b>Remerciements .....</b>	<b>2</b>
<b>INTRODUCTION BIBLIOGRAPHIQUE .....</b>	<b>8</b>
<b>Chapitre 1 : Les rétrovirus .....</b>	<b>9</b>
<b>I- Caractéristiques générales des rétrovirus .....</b>	<b>9</b>
A. Structure de la particule virale .....	9
B. L'organisation génomique des rétrovirus .....	11
C. Les protéines régulatrices et accessoires .....	13
<b>II- Classification des rétrovirus .....</b>	<b>14</b>
<b>III- Endogénéisation des rétrovirus .....</b>	<b>16</b>
A. Généralités .....	16
B. Rôles des ERVs dans la physiologie de leurs hôtes .....	16
C. Risques potentiels liés à la présence de rétrovirus endogènes .....	17
<b>IV- Le cycle de réplication .....</b>	<b>18</b>
A. La phase précoce .....	18
B. La phase tardive .....	19
<b>V- Quelques données sur les rétrovirus ASLVs et PERVs .....</b>	<b>21</b>
A. Les rétrovirus aviaires de la famille des ASLVs .....	21
B. Les rétrovirus endogènes porcins .....	24
<b>Chapitre II : L'intégrase et l'intégration rétrovirale .....</b>	<b>28</b>
<b>I- Le transport nucléaire et le complexe de préintégration (PIC) .....</b>	<b>28</b>
<b>II- Le mécanisme d'intégration .....</b>	<b>30</b>
A. Clivage en 3' et transfert de brins .....	31
B. Désintégration .....	32
C. Réparation .....	32
D. L'intégration concertée .....	34
<b>III- L'intégrase .....</b>	<b>35</b>
A. Généralités sur l'intégrase .....	35
B. Oligomérisation de l'IN .....	38
C. Structure cristallographique de l'IN .....	40
D. Les inhibiteurs de l'IN .....	42
<b>IV- Les sites d'intégration .....</b>	<b>44</b>
A. Les profils d'intégration .....	44
B. Les paramètres de sélection .....	45
C. La sélection des sites d'intégration et la thérapie génique .....	47
<b>Chapitre III : Les cofacteurs cellulaires de l'intégration .....</b>	<b>48</b>
<b>I- Les constituants cellulaires du PIC .....</b>	<b>48</b>
A. HMGA1a .....	48

B. BAF .....	49
<b>II- Les partenaires cellulaires de l'IN de VIH-1 .....</b>	<b>49</b>
A. Les cofacteurs impliqués dans les étapes pré-intégratives .....	49
B. Les cofacteurs impliqués dans l'intégration .....	51
C. Les cofacteurs impliqués dans les étapes post-intégratives.....	55
<b>III- Les partenaires cellulaires de l'IN de MLV .....</b>	<b>58</b>
A. Les facteurs de réparation de l'ADN .....	58
B. Les facteurs de liaison à l'ARN .....	58
C. Les facteurs de liaison à la chromatine .....	59
D. Brd2, partenaire potentiel de l'IN de MLV et de VIH-1.....	60
<b>OBJECTIFS.....</b>	<b>64</b>
<b>MATERIELS ET METHODES .....</b>	<b>67</b>
<b>I- Construction des vecteurs d'expression de l'IN PERV et ses domaines .....</b>	<b>68</b>
A. PCR.....	68
B. Digestion et ligation .....	68
D. Séquençage.....	69
<b>II- Production des intégrases d'ASLV et de PERV .....</b>	<b>71</b>
A. Production en bactérie.....	71
B. Production en baculovirus/cellule d'insecte.....	71
<b>III- Purification des intégrases .....</b>	<b>72</b>
A. Les solutions .....	72
B. Le protocole.....	73
<b>IV- Analyse et dosage des fractions d'IN purifiées sur gel SDS-PAGE .....</b>	<b>73</b>
<b>V- Analyse des formes oligomériques de l'intégrase par fixation chimique .....</b>	<b>74</b>
<b>VI- Tests des activités catalytiques de l'intégrase.....</b>	<b>75</b>
A. Principe .....	75
B. Protocole.....	76
<b>VII- Test par Far Western Blot .....</b>	<b>77</b>
A. Principe .....	77
B. Les étapes critiques d'une analyse par far western blot .....	78
C. Protocole.....	79
<b>RESULTATS .....</b>	<b>81</b>
<b>I- Objectifs .....</b>	<b>82</b>
<b>II- Article .....</b>	<b>84</b>
<b>III- Données supplémentaires.....</b>	<b>113</b>
A. Purification de l'IN d'ASLV.....	113
B. Cross link de l'IN d'ASLV .....	114

C. Activités catalytiques de l'IN d'ASLV .....	114
D. Test de l'intégration concertée de PERV .....	115
E. Interaction INs-TAB182.....	117
<b>IV- Discussion des résultats .....</b>	<b>119</b>
A. Comparaison des rendements de protéines pures .....	119
B. Activités catalytiques des intégrases .....	120
C. Spécificité de l'interaction IN PERV/ Brd2 .....	122
D. Caractérisation de l'interaction entre les domaines de l'IN PERV et Brd2.....	122
<b><i>CONCLUSION GENERALE ET PERSPECTIVES.....</i></b>	<b><i>124</i></b>
<b>I- Conclusion générale.....</b>	<b>125</b>
<b>II- Perspectives expérimentales.....</b>	<b>125</b>
B. Validation de l'interaction IN PERV-Brd2 .....	126
C. Evaluation du rôle de Brd2 dans l'intégration .....	127
<b>III- Perspectives à long terme.....</b>	<b>130</b>
A. Thérapies antirétrovirales.....	130
B. Prévention contre l'intégration du PERV et xénotransplantation .....	130
<b><i>REFERENCES .....</i></b>	<b><i>132</i></b>

## Liste des abréviations

<b>ALV:</b> Avian leukosis virus	<b>ODN:</b> oligonucléotides double-brin
<b>ARN:</b> Acide ribonucléique	<b>PBS:</b> Primer binding site
<b>ARNm:</b> ARN messenger	<b>PERV:</b> Porcine endogenous retrovirus
<b>ARNt:</b> ARN de transfert	<b>PIC:</b> Pre-integration complex
<b>BAF:</b> Barrier to autointegration factor	<b>PPT:</b> Polypurine tract
<b>Brd2:</b> Bromodomain containing 2 protein	<b>PR:</b> Protéase
<b>CA:</b> Capside	<b>Psi:</b> Packaging signal
<b>CCD:</b> Catalytic Central Domain	<b>RRE:</b> Rev response element
<b>CpG:</b> Dinucléotide CG nom méthylé	<b>RSV:</b> Rous sarcoma virus
<b>CTD:</b> Carboxy-terminal domain	<b>RT:</b> Rétrotranscriptase
<b>DNase I:</b> Deoxyribonuclease I	<b>RTC:</b> Reverse transcription complex
<b>EED:</b> Embryonic Ectoderm Development	<b>SA:</b> Site accepteur de l'ARN viral
<b>EOPS:</b> Exempts d'organismes pathogènes spécifiques	<b>SAXS:</b> Small-angle X-ray
<b>FeLV:</b> Feline leukemia virus	<b>SD:</b> Site donneur
<b>FIV:</b> Feline immunodeficiency virus	<b>U3:</b> Unique sequence element at the 3' end
<b>HEK-293:</b> Human embryonic kidney cells	<b>U5:</b> Unique sequence element at the 5' end
<b>HERV:</b> Human endogenous retrovirus	<b>UTR:</b> Untranslated region
<b>HMGA1a:</b> High mobility group chromosomal protein A1	<b>VIH:</b> Human immunodeficiency virus
<b>Hsp60:</b> Heat shock protein60	<b>VBPI:</b> Von Hippel Lindau binding protein 1
<b>HTLV-1:</b> Human T-lymphotropic virus 1	<b>Vif:</b> Viral Infectivity Factor
<b>IN:</b> Intégrase	<b>Vpr:</b> Viral protein R
<b>INI1:</b> Integrase interactor-1	<b>Vpu:</b> Viral protein U
<b>IRES:</b> Internal ribosome entry site	
<b>JSRV:</b> Jaagsiekte sheep retrovirus	
<b>kb:</b> Kilobase	
<b>KoRV:</b> Koala retrovirus	
<b>LEDGF/p75:</b> Lens epithelium growth factor	
<b>LTR:</b> Long terminal repeat	
<b>MA:</b> Matrice	
<b>MLV:</b> Murine leukemia virus	
<b>MMTV:</b> Mouse mammary tumor virus	
<b>NC:</b> Nucléocapside	
<b>nt:</b> Nucléotide	
<b>NTD:</b> amino-terminal domain	

## **INTRODUCTION BIBLIOGRAPHIQUE**

## Chapitre 1 : Les rétrovirus

Les rétrovirus sont des virus ARNs, enveloppés qui présentent deux particularités : (i) la réverse transcription de l'ARN viral en ADN, réalisée par une enzyme virale, la transcriptase inverse (TI ou RT pour reverse transcriptase), et (ii) l'intégration de cet ADN viral dans le génome de la cellule infectée grâce à une enzyme virale appelée l'intégrase (IN).

Le but de ce premier chapitre est d'introduire les deux modèles de rétrovirus étudiés : les virus de sarcomes et de leucoses aviaires (ASLV) et les rétrovirus endogènes porcins (PERV). Ces deux familles de rétrovirus présentent des caractéristiques communes mais aussi des divergences dues à leur situation dans la famille virale. Afin de compléter ces modèles, j'ai utilisé des données bibliographiques correspondant à un autre virus du même genre que le virus PERV mais plus étudié, le virus de la leucémie murine (MLV), et à un lentivirus largement étudié, le virus de l'immunodéficience humaine (VIH-1).

### **I- Caractéristiques générales des rétrovirus**

#### *A. Structure de la particule virale*

Les rétrovirus sont des virus à ARN, enveloppés, de diamètre compris entre 80 et 120 nanomètres. Ils sont constitués de protéines structurales, de protéines enzymatiques, et, pour certains d'entre eux, de protéines régulatrices et accessoires. Quant au génome viral, celui-ci est formé de deux brins monocaténares d'ARN de polarité (+), comportant 7000 à 12000 nucléotides, reliés par des ponts hydrogènes à leur extrémité 5' (figure 1).

#### *a) L'enveloppe virale*

L'enveloppe externe est constituée d'une bicouche lipidique issue de la dernière cellule infectée par bourgeonnement, et de la glycoprotéine de surface (SU) et la protéine transmembranaire (TM). Ces deux protéines ont un rôle déterminant dans la fixation du virus sur un récepteur spécifique de la membrane cellulaire et la fusion des membranes virales et cellulaires lors de la pénétration dans la cellule.

#### *b) Les protéines de structure*

- La matrice (MA), 15-20 kDa selon les rétrovirus. Elle tapisse la face intérieure de l'enveloppe virale en interagissant avec la membrane lipidique interne grâce à son

groupement myristate. Elle possède une séquence de localisation nucléaire (NLS), et participe au sein complexe de préintégration (PIC) avec d'autres protéines virales et cellulaires, à l'adressage de l'ADN viral au noyau (Haffar OK, Popov S et al. 2000; Freed EO 2001).

- La capsid (CA), 24-30 kDa, se caractérise par la présence d'une région majeure d'homologie MHR. Elle possède une forme icosaédrique et renferme le génome viral et des enzymes (Freed EO 2001).
- La nucléocapsid (NC), protéine de 10-15k Da, avec des domaines de doigt de zinc type CCHH (Cys-His-Cys-X<sub>2</sub>-Cys-X<sub>4</sub>-His-X<sub>4</sub>-Cys) (Covey SN 1986) flanquées par des séquences basiques impliquées dans la liaison avec les acides nucléiques (Cimarelli A, Sandin S et al. 2000). Les structures en doigt de zinc contribuent à l'encapsidation du génome viral par son interaction avec la séquence Psi (Freed EO 2001; Muriaux D and Darlix JL 2010).

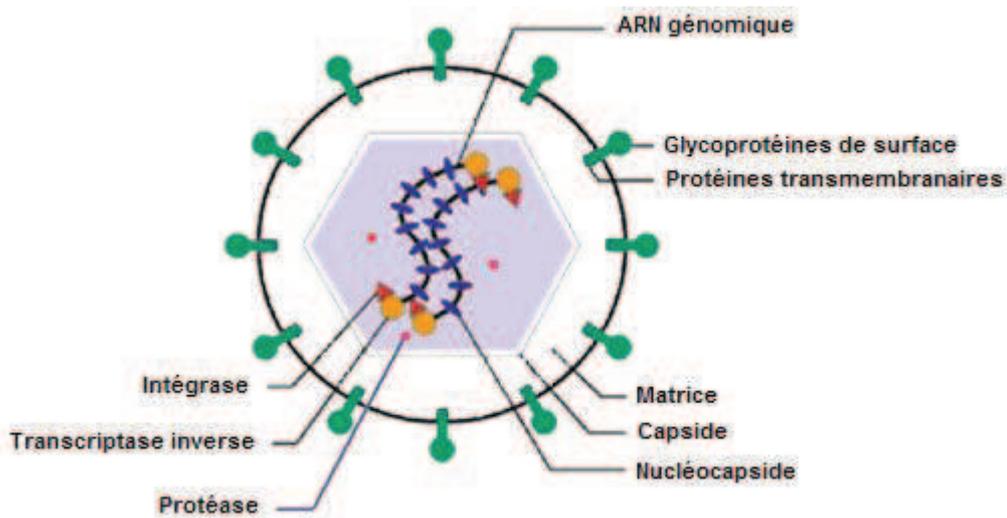
### *c) Les enzymes virales*

- La réverse transcriptase (RT), 50-95 kDa ; elle catalyse la synthèse d'ADN à partir du génome d'ARN viral (baltimore, 1970). Il s'agit d'un hétérodimère constitué de deux sous-unités douées de 2 activités principales : une activité ADN polymérase capable d'incorporer les déoxyribonucléotides et une activité RNaseH ARN/ADN dépendante. L'activité ADN polymérase de la RT est généralement lente (1 à 100 nucléotides/seconde) et présente une faible fidélité (10<sup>-4</sup> erreurs/base incorporée). Contrairement à l'ADN polymérase de l'hôte, elle est incapable de corriger les bases erronées. Par conséquent, au moins une mutation/génome et par cycle de reverse transcription est introduite (fields of virology, fifth edition 2007). Notons que cette incapacité d'édition de la RT est à la base de la forte variabilité observée chez le VIH-1.

La RNase H est une exonucélase capable de dégrader uniquement l'ARN sous forme hybride (ARN : ADN) dans les deux sens 5' → 3' et 3' → 5'. Cette action engendre les oligonucléotides 3'OH et 5'-P de 6 à 10 résidus de long, qui serviront d'amorces pour l'initiation de la synthèse d'ADN grâce à l'activité ADN polymérase de la RT (Frankel AD and Young JA 1998; Freed EO 2001).

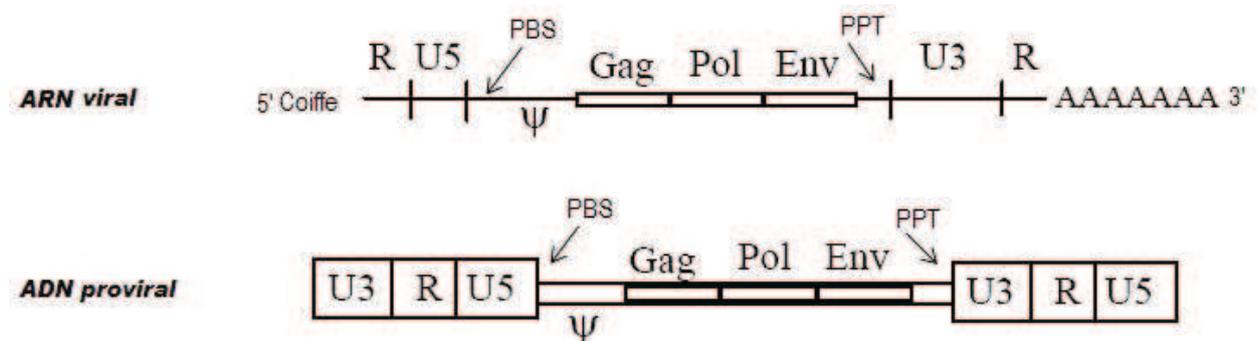
- L'intégrase (IN), 32-46 kDa, permet l'intégration de l'ADN bicaténaire viral dans le génome cellulaire, elle sera détaillée dans le chapitre II.

- La protéase (PR), 15 kDa, est une petite enzyme (100 aa) de la famille des aspartyl-protéases. Le premier clivage catalysé par PR se produit durant ou immédiatement après la libération du virion. Ce clivage est vraisemblablement intramoléculaire (pettit et al., 2004) et permet la libération de PR de son précurseur Gag-Pol. La protéine active est sous forme d'homodimère. Elle clive spécifiquement les liaisons Phe-Pro et Tyr-Pro, qui ne sont pas hydrolysées par les protéases des mammifères. Ce clivage induit la maturation des précurseurs Gag et Gag-Pro-Pol qui participe au changement de la morphologie des virions et conduit à leur maturation (Frankel AD and Young JA 1998; Freed EO 2001).



**Figure 1. Représentation d'un rétrovirus.** La particule virale est composée (i) d'une enveloppe constituée de glycoprotéines de surface, et de protéines transmembranaires, (ii) de protéines de structure (Matrice, Capside et Nucléocapside), (iii) d'enzymes (Intégrase, Transcriptase inverse et Protéase) et (iv) d'un génome, formé de 2 brins monocaténares d'ARN de polarité + (Voisset C and Andrawiss M 2000).

### B. L'organisation génomique des rétrovirus



**Figure 2. Génome rétroviral.** L'ARN viral comporte en 5' une coiffe, suivie d'une petite région R (18-250 nt), répétée à chacune des extrémités. Cette région contient le signal de polyadénylation (AAUAAA). La séquence PBS (Primer Binding Site) est une séquence qui correspond à 18 nt qui fixe spécifiquement un ARNt cellulaire et sert d'amorce pour l'initiation de la transcription inverse. Le signal d'encapsidation  $\psi$  va permettre

l'incorporation de l'ARN génomique dans les particules virales néo-synthétisées. Les régions Gag, Pol et Env consistent les 3 gènes essentiels responsables de l'expression des protéines structurales et enzymatiques. La séquence PPT (PolyPurine Tract) est région riche en A et G ; elle comporte 10-12 nt et est responsable de l'initialisation du brin d'ADN(+) pendant la transcription inverse. Le génome viral se termine en 3' par une queue poly A. U3 est une région unique non codante (200-1200nt). Elle formera l'extrémité 5' de l'ADNc et comporte le promoteur viral. La transcription inverse de l'ARN viral par la RT conduit à un ADN double brin. La duplication des segments terminaux va créer les LTRs (U3-R-U5). Cet ADN sera intégré dans le génome de la cellule hôte, et constitue l'ADN proviral.

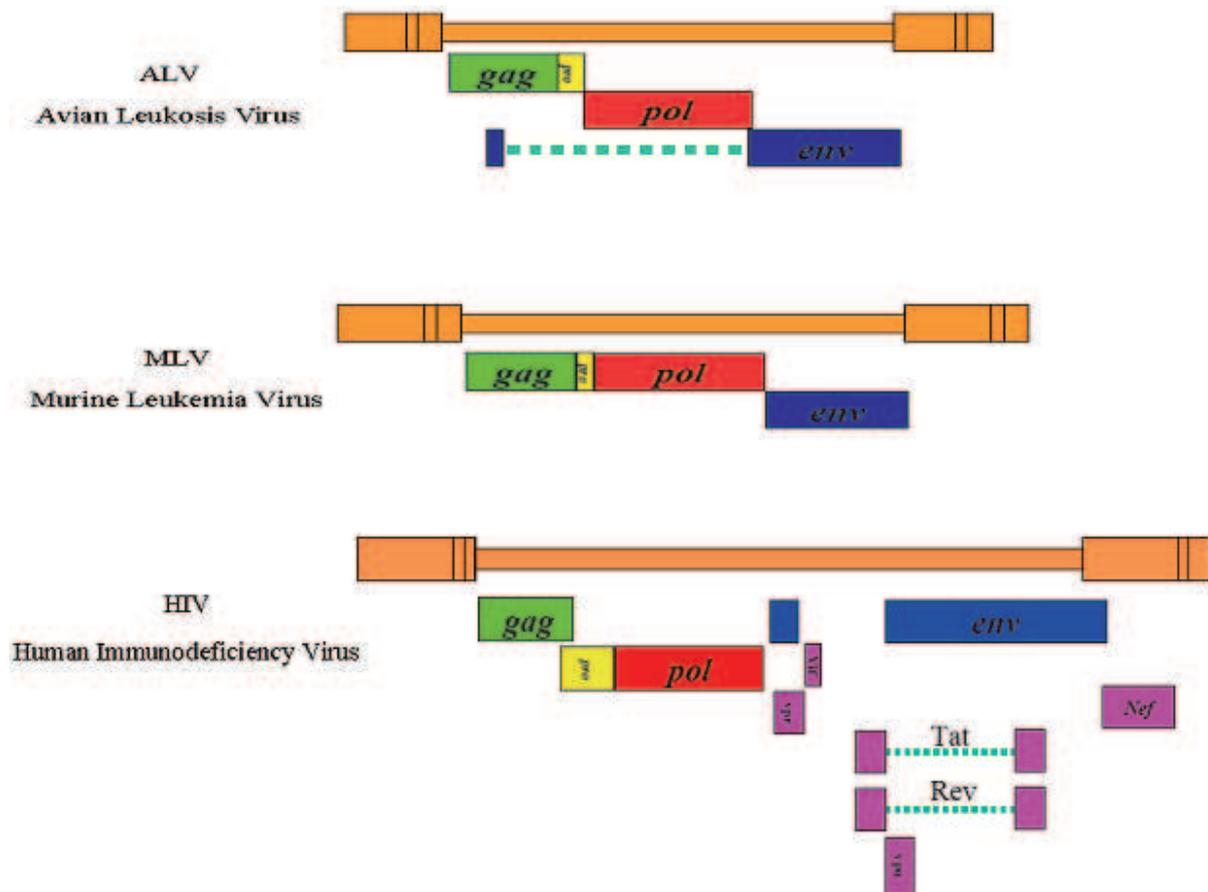
Le génome rétroviral est constitué de différentes régions, ayant chacune un rôle bien défini. Il contient, au minimum, 3 gènes communs à tous les rétrovirus : *gag* (« group specific antigen »), *pol* (« polymérase ») et *env* (« enveloppe »). Aux extrémités du génome se trouvent des séquences terminales non codantes, les LTRs (Long Terminal Repeats) (figure 2). Les virus de type ASLVs et PERVs possèdent un génome simple.

- Les LTRs : Chaque LTR se subdivise en séquences U3 (Unique en 3'), R (Redondante) et U5 (Unique en 5'). Bien qu'elles soient identiques, ces séquences ne jouent pas le même rôle : en 5', le LTR contient le promoteur régulant l'expression du génome viral et en 3', il apporte le signal de fin de transcription.

- Le gène *gag* code pour un précurseur polypeptidique Gag myristylé dont le clivage enzymatique par la protéase produit les trois protéines de structure, la capsid (CA), la matrice (MA) et la nucléocapsid (NC), et d'autres protéines comme p12 chez MLV (Prizan-Ravid A, Elis E et al. 2010).

- Les gènes *pro* et *pol* (« polymérase ») sont exprimés sous la forme d'un précurseur polypeptidique Gag-Pro-Pol générant la transcriptase inverse, l'intégrase et la protéase ; sachant que l'enzyme protéase peut être codée soit par le gène *gag*, soit par le gène *pol* selon le type de virus. La protéase est libérée par un mécanisme d'autocatalyse, alors que la transcriptase inverse et l'intégrase sont clivées par la protéase virale.

- Le gène *env* code un précurseur polyprotéique Env. Son clivage par une protéase cellulaire permet la production d'une protéine de surface et d'une protéine transmembranaire. Cette dernière porte le peptide fusogène R effecteur de la fusion.



**Figure 3. Exemples d'organisation des génomes de rétrovirus simples et complexes.** Les gènes *gag*, *pro*, *pol* et *env* sont représentés respectivement par des cadres de couleur vert, jaune, rouge et vert. Les gènes supplémentaires des rétrovirus complexes sont représentés en violet.

(D'après <http://www.microbiologybytes.com/virology/Retroviruses.html>).

### C. Les protéines régulatrices et accessoires

Les rétrovirus à génome complexe possèdent également des protéines régulatrices et des protéines accessoires codées par des gènes auxiliaires. (pour une revue générale sur les protéines régulatrices et accessoires de HIV-1, voir (Seelamgari A, Maddukuri A et al. 2004)). Parmi les plus importantes, on retrouve les protéines régulatrices Tat (transactivator of transcription) et Rev (Regulation of Expression of Viral proteins). Tat, en recrutant l'ARN polymérase cellulaire, constitue un activateur puissant de la transcription de l'ARN viral (Johri MK, Mishra R et al. 2011); Rev est une protéine primordiale pour l'export nucléaire des ARNm néosynthétisés, incomplètement épissés, codant pour les protéines structurales du virus. Afin de traverser les pores nucléaires, Rev se lie à la séquence RRE (Rev-Responsive Element) des ARNm et recrute l'exportine1 et Ran (Grewe B and Uberla K 2010).

Le VIH-1 code pour 4 protéines accessoires (Li L, Li HS et al. 2005).

- Nef (Negative regulatory Factor) est une protéine de 27 kDa qui permet de maintenir l'infectivité du virus en perturbant différentes fonctions immunologiques de la cellule hôte. Elle provoque la diminution de l'expression des récepteurs CD4 et CMH-1 impliqués dans la reconnaissance des cellules infectées par le système immunitaire. Elle est également capable de se lier à plusieurs protéines de transduction du signal cellulaire, augmentant ainsi la production virale.

- Vif (Viral Infectivity Factor) est une protéine de 23 kDa qui désamorce un système efficace de défense cellulaire en dégradant la protéine cellulaire APOBEC3G, capable d'induire des mutations dans le génome viral qui sont délétères pour le virus.

- Vpr (Viral Protein R) est une protéine basique de 14 kDa qui confère aux lentivirus l'avantage d'infecter des cellules non multipliantes (monocytes et macrophages), en participant à la translocation cytoplasme-noyau du PIC. Indépendamment de tout contexte viral, il a été montré que VPR interfère avec les processus de multiplication cellulaire en bloquant le cycle en G2.

- Vpu (Viral Protein U) possède deux fonctions principales chez le VIH. La première est de séquestrer le récepteur viral CD4 à l'intérieur de la cellule infectée afin de le dégrader, empêchant la surinfection et facilitant le bourgeonnement des produits viraux. La seconde est de bloquer l'action de la protéine cellulaire Tetherin (BST-2, CD317, HM1.24), qui empêche leur bourgeonnement (Dubé M, Bego MG et al. 2010).

## **II- Classification des rétrovirus**

Actuellement, la famille des rétrovirus (*Retroviridae*) a été divisée en sept genres par l'ICTV (International Committee on Taxonomy of Viruses) (<http://ictvonline.org>). Cette classification est basée essentiellement sur la comparaison des séquences du gène *pol*. On distingue les Alpharétrovirus, les Betarétrovirus, les Gammarétrovirus, les Deltarétrovirus, les Epsilon-rétrovirus, les Lentivirus et les Spumavirus (figure 4).

- Les Alpharétrovirus sont trouvés chez les oiseaux. On distingue les virus leucémogènes (ALV) tel le virus RAV-1 (Rous Associated Virus type 1), et les virus sarcomatogènes (ASV) tel le virus RSV. Ils peuvent induire des sarcomes et autres tumeurs.

- Les Betarétrovirus (exemple le Jaagsiekte sheep retrovirus).

- Les Gammarétrovirus sont responsables de sarcomes et de leucémies. On y trouve le Murine leukemia virus et le Porcine endogenous retrovirus.

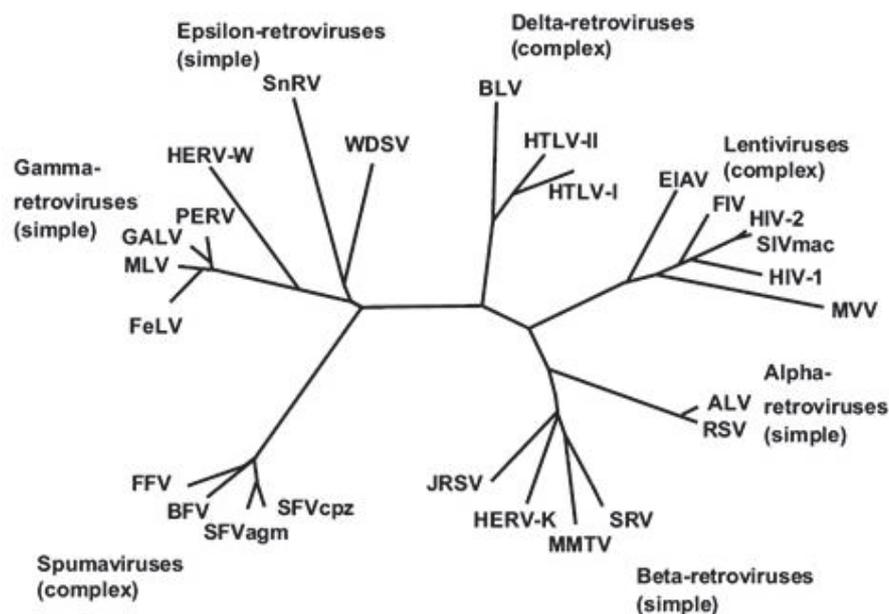
- Les Deltarétrovirus sont responsables de leucémies (exemple le Human T-lymphotropic virus Type I).

- Les Epsilonrétrovirus sont trouvés chez les poissons (exemple Walleye dermal sarcoma virus).

- Les Lentivirus, dont le VIH-1. Ils sont caractérisés par une longue période d'incubation, d'où leur nom lentivirus qui vient du latin *lentis* signifiant lent. Ils sont tous cytopathogènes et sont responsables d'atteintes du système nerveux, des poumons, des systèmes articulaires et des cellules hématopoïétiques.

- Les Spumavirus se caractérisent par des infections chroniques sans apparition de signes cliniques. Leur nom vient du latin (*spuma* = écume) dû au fait que l'infection, *in vitro*, se caractérise par l'apparition de syncytium et de vacuoles donnant aux cellules un aspect spumeux.

Se basant sur la constitution du génome, les sept genres rétroviraux peuvent être regroupés en deux grandes catégories : les rétrovirus à génome simple représentés par les alpha, beta, gamma et epsilon rétrovirus, et ceux à génome complexe représentés par les lentivirus, les deltavirus et les spumavirus.



**Figure 4. Phylogénie des rétrovirus.** Adapté de Weiss (Weiss RA 2006); les rétrovirus sont divisés en 7 genres par l'ICTV. Les genres Alpha-, Beta-, Gamma-, Epsilonrétrovirus et lentivirus comportent des rétrovirus endogènes.

### III- Endogénéisation des rétrovirus

#### *A. Généralités*

L'endogénéisation est un mécanisme par lequel un rétrovirus infecte une cellule appartenant à la lignée germinale d'un individu, et sera transmis à la descendance selon la loi Mendélienne (transmission verticale). À partir de ce moment, toutes les cellules du nouvel organisme posséderont une copie du génome viral dans leur ADN. Le provirus intégré est dit « endogène » (pour une revue détaillée sur l'origine et l'évolution des ERVs voir (Ribet D and Heidmann T 2010)).

On estime que chez l'homme, le génome contient 8% de séquences d'origine rétrovirale (Lander ES, Linton LM et al. 2001; Mayer J and Meese E 2005). Jusqu'à présent, on considérait que seuls les rétrovirus à génomes simples présentaient des formes endogènes. Mais cela est remis en question depuis la découverte récente d'un lentivirus endogène dans le génome de lapin européen, le rabbit endogenous lentivirus type K (RELK) (Katzourakis A, Tristem M et al. 2007).

Il existe un exemple de rétrovirus actuellement en cours d'endogénéisation chez un mammifère : le virus KoRV (Koala retrovirus) (Tarlinton RE, Meers J et al. 2006). Il constitue de ce fait un outil très intéressant pour étudier les modifications du virus initial aboutissant à son endogénéisation et les conséquences de ce mécanisme pour l'organisme hôte.

#### *B. Rôles des ERVs dans la physiologie de leurs hôtes*

Il est admis que certains ERVs peuvent avoir des fonctions biologiques et peuvent être bénéfiques pour l'hôte.

On citera :

- Un rôle dans la régulation des gènes :

Les LTRs des séquences rétrovirales endogènes contiennent diverses séquences promotrices et régulatrices qui peuvent moduler l'expression des gènes cellulaires. Par exemple, le LTR de la famille HERV-E (Human Endogenous Retrovirus E) semble être impliqué dans la régulation de l'expression de l'Apolipoprotéine C1 dans le foie, du récepteur de l'endothéline B dans le placenta et du facteur de croissance insuline-like 4 (INSL4) (Cohen CJ, Lock WM et al. 2009).

- Un rôle dans la placentation :

C'est le cas du rétrovirus endogène humain HERV-W. Il contient un locus unique nommé ERVWE1 codant pour une protéine d'enveloppe complète appelée Syncytine. Cette protéine virale, codée par le gène *env* durant certaines étapes de la grossesse, est impliquée dans la formation du placenta, plus précisément dans le mécanisme de fusion des trophoblastes (Mi S, Lee X et al. 2000). La syncytine contient également une région immunosuppressive qui pourrait aussi jouer un rôle dans la protection du fœtus face à la réponse immunitaire maternelle.

- Un rôle dans la protection contre les rétrovirus exogènes:

Chez les poulets, l'expression de protéines virales endogènes aboutit à une « tolérance » du système immunitaire vis-à-vis de ce type de protéines. Il en résulte une réponse immunitaire moins exacerbée (mais néanmoins protectrice) lors de l'infection par un virus exogène (Crittenden LB and Fadly AM 1985; Kuhnlein U, Fairfull RW et al. 1993). Chez les souris, l'expression des protéines d'enveloppe et de capsid par les loci Fv-1 et Fv-4, respectivement, conduit à une résistance de la cellule à l'infection par un virus exogène (Nethe M, Berkhout B et al. 2005; Takeda A and Matano T 2007 ).

### *C. Risques potentiels liés à la présence de rétrovirus endogènes*

Un risque lié aux ERVs concerne tous les produits issus de cultures cellulaires, en particulier les préparations des vaccins, des vecteurs rétroviraux destinés à la thérapie génique, et les tissus ou organes animaux dans le cadre des xéno greffes (Dewannieux M, Ribet D et al. 2010). Dans ces préparations, les particules rétrovirales endogènes doivent être inactivées ou éliminées. L'expression des ERVs a été tout de même rapportée dans diverses pathologies, sans que l'on sache clairement quelles sont les conséquences de cette expression (Lachuer J, Ronfort C et al. 1996; Roelofs H, van Gurp RJ et al. 1998; Wang-Johanning F, Frost AR et al. 2001; Wang-Johanning F, Frost AR et al. 2003; Wang-Johanning F, Liu J et al. 2007; Balada E, Vilardell-Tarrés M et al. 2010; Antony JM, Deslauriers AM et al. 2011).

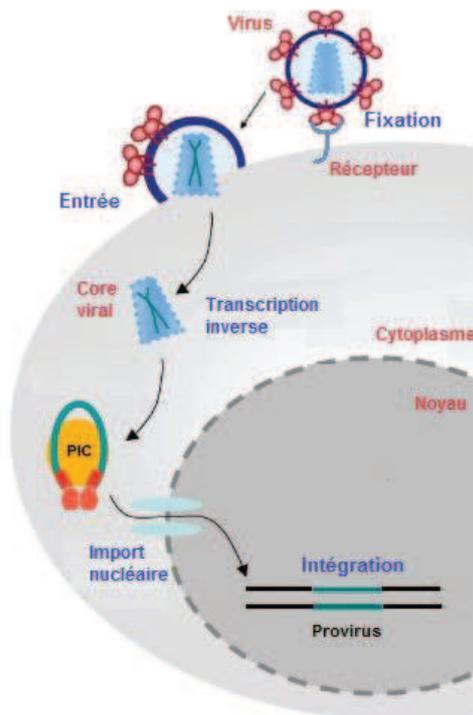
## IV-Le cycle de réplication

Le cycle répliatif des rétrovirus peut se diviser en deux phases principales : la phase pré-intégrative (ou précoce) et la phase post-intégrative (ou tardive).

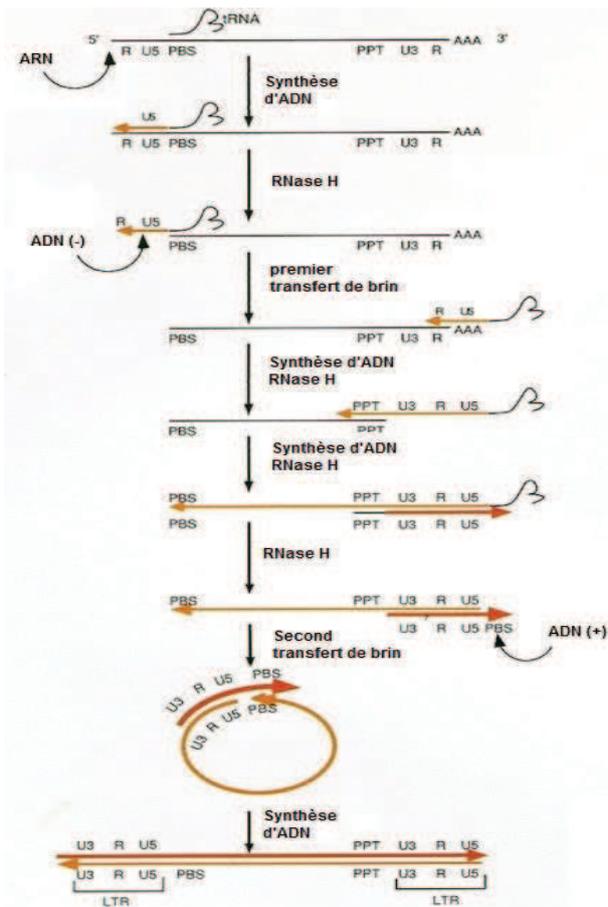
### *A. La phase précoce (figure 5)*

Le cycle de réplication des rétrovirus est initié par une interaction spécifique entre le récepteur cellulaire et la glycoprotéine de surface (SU). Cela aboutit à un changement de conformation du complexe SU-TM permettant le démasquage du peptide de fusion porté par la TM. Il en résulte une fusion de l'enveloppe virale avec la membrane cellulaire. Le complexe de réverse transcription (RTC) est alors libéré dans le cytoplasme de la cellule hôte. Sous l'action de la transcriptase inverse virale, l'ARN rétroviral simple brin subit une transcription inverse à l'issue de laquelle est formé un ADN double brin, complémentaire (ou ADNc) (figure 6).

A ce stade, l'ADN se trouve dans le complexe de préintégration (PIC) en association avec des protéines virales et cellulaires. Ce complexe est ensuite transporté dans le noyau de la cellule hôte. Une fois dans le noyau, l'ADN viral est intégré par le biais de l'intégrase dans le génome de la cellule hôte sous forme de provirus. Les étapes comprises entre la formation du PIC et l'intégration seront détaillées dans le chapitre suivant.



**Figure 5. La phase précoce du cycle de réplication rétrovirale.** La phase précoce débute par la reconnaissance spécifique d'un ou plusieurs récepteurs cellulaires par les glycoprotéines d'enveloppe virales exposées à la surface des virions, conduisant à l'entrée virale, et s'achève par l'intégration des provirus au sein du génome cellulaire. Notons que l'étape d'import nucléaire, telle qu'elle est représentée sur ce schéma, est caractéristique des lentivirus. Pour MLV, PERV et ASLV, l'association du PIC avec le noyau de la cellule hôte dépend dans certains cas de la rupture de la membrane nucléaire lors de la mitose. L'ADN viral et l'ADN cellulaire sont représentés respectivement par des traits de couleur vert et noir. (Adapté du : [http://cochin.inserm.fr/la\\_recherche/departements/bio\\_cihp/equipe-berlioz-torrent-emiliani/les-etapes-nucleaires-de-la-replication-du-vih-1-1](http://cochin.inserm.fr/la_recherche/departements/bio_cihp/equipe-berlioz-torrent-emiliani/les-etapes-nucleaires-de-la-replication-du-vih-1-1))



**Figure 6. Le processus de réverse transcription (Coffin JM, Hughes SH et al. 1997).** La réverse transcription est initiée par un recrutement d'un ARNt au niveau de la séquence PBS de l'ARN viral. Il y a donc synthèse d'un court brin d'ADN noté ADN(-) en 5', qui sera ensuite transféré à l'extrémité 3' de l'ARN via l'homologie de séquence entre les régions R. L'élongation se poursuit vers l'extrémité 5' formant ainsi le premier brin d'ADN. En parallèle, la fonction RNase de l'enzyme transcriptase inverse dégrade progressivement la matrice ARN, à l'exception des séquences PPT et cPPT, qui vont permettre l'initiation de la synthèse d'un second brin, l'ADN(+). L'élongation de ce dernier jusqu'en 3' est suivie d'une dissociation de l'ARNt. L'homologie des séquences des PBS permet au brin (+) d'être transféré vers le (-) pour finir sa formation. Ce processus se termine par une synthèse simultanée des LTRs sur les 2 brins d'ADN entraînant la formation d'une copie de séquences U3-R-U5.

### B. La phase tardive (figure 7)

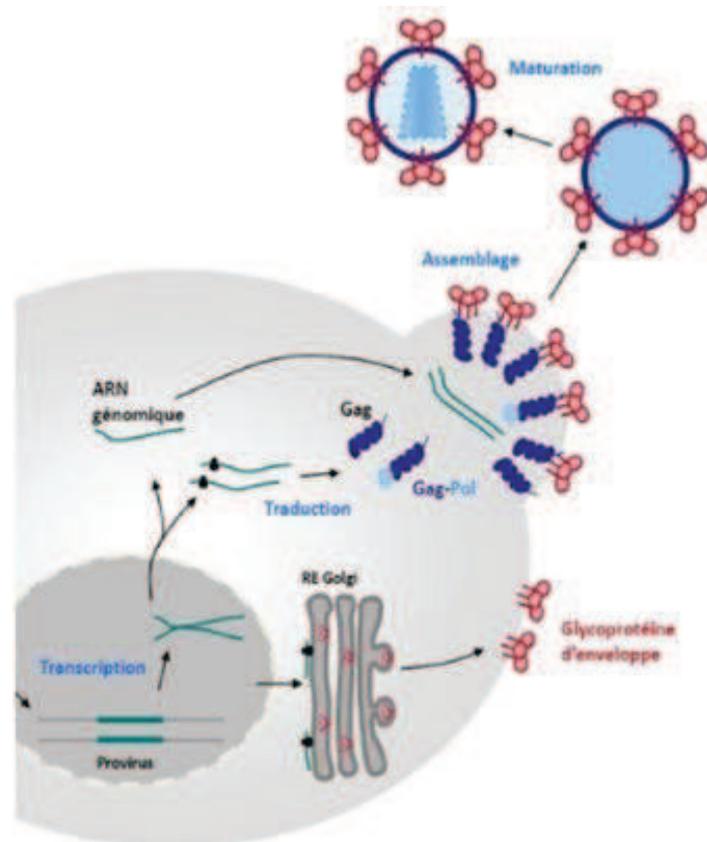
Cette phase débute par la transcription des gènes viraux. Le provirus est transcrit par l'ARN polymérase II cellulaire en un seul transcrit primaire (ARN génomique) à partir du premier nucléotide de R du LTR 5'. Le promoteur et les séquences activatrices (*enhancers*) de la transcription sont situés dans la région U3. Le transcrit représente le génome viral complet R-U5-gag-pol-env-U3-R, auquel s'ajoutent une coiffe en 5' et le poly A (environ 150 nucléotides) en 3'. Une partie des transcrits viraux va subir des processus d'épissage, alors que l'ARNm non épissé va coder les protéines Gag et Pol et/ou sera dimérisé et encapsidé dans de nouvelles particules en formation.

Chez les rétrovirus complexes tel le VIH-1, le transport dans le cytoplasme de l'ARN génomique est réalisé grâce une interaction Rev/RRE. Les ARN viraux comportent une coiffe en 5' mais également des séquences IRES (internal ribosome entry site) d'entrée interne du ribosome. Ces séquences sont impliquées dans la traduction des protéines virales (Berlioz C and Darlix JL 1995; Deffaud C and Darlix JL 2000; Brasey A, Lopez-Lastra M et al. 2003; Ronfort C, De Breyne S et al. 2004 ). Les précurseurs Gag/Pol et Pol sont synthétisés à partir

de l'ARNm génomique par les ribosomes « libres ». Le précurseur Env est traduit par les ribosomes liés à la membrane du réticulum endoplasmique.

L'étape d'assemblage s'opère à la membrane plasmique de la cellule, ou peut avoir lieu au niveau des membranes intracellulaires et les virions seront ensuite sécrétés par une voie de sécrétion endosomale vers la membrane plasmique (Basyuk E, Galli T et al. 2003; Gould SJ, Booth AM et al. 2003). Le précurseur Gag est au cœur du processus d'assemblage (Muriaux D, Costes KS et al. 2004; Muriaux D and Darlix JL 2010).

Enfin, le bourgeonnement conduit au relâchement de la nouvelle particule virale en dehors de la cellule. La maturation des précurseurs polyprotéiques Gag et Gag-Pol a lieu pendant ou après le bourgeonnement. Les MA, CA et NC se réarrangent ensuite pour donner naissance à une nouvelle particule virale infectieuse.



**Figure 7. La phase tardive du cycle de réplication rétrovirale.** Les étapes tardives correspondent à la synthèse, l'assemblage des composants viraux avec le génome viral, et la production de particules virales matures infectieuses. (D'après :[http://cochin.inserm.fr/la\\_recherche/departements/biocihp/equipe-berlioz-torrent-emiliani/les-etapes-tardives-de-la-replication-virale-1](http://cochin.inserm.fr/la_recherche/departements/biocihp/equipe-berlioz-torrent-emiliani/les-etapes-tardives-de-la-replication-virale-1))

## **V- Quelques données sur les rétrovirus ASLVs et PERVs**

### *A. Les rétrovirus aviaires de la famille des ASLVs*

#### *a) Généralités sur les ASLVs*

Les virus de leucoses et de sarcomes aviaires (ASLVs) sont classés dans le genre alpharetrovirus. Décrits depuis le début du siècle dernier, ils comprennent un nombre important de souches et d'isolats viraux issus du poulet. On distingue :

i) Les virus réplicatifs et dépourvus de gènes oncogènes, comme les virus de leucoses aviaires (ALVs). L'inoculation d'un ALV à des poussins induit l'apparition de leucoses, mais celles-ci ne surviennent qu'après un délai important. La transformation est liée à l'intégration et à l'activation par les LTRs viraux d'un oncogène (*myc*, *erbB*) dans les cellules cibles du virus. Le fait que les ALVs se transmettent verticalement de la poule aux poussins a permis l'élaboration de techniques permettant de détecter la protéine de la CA (p27) des ALVs dans l'albumen de l'œuf, et d'éliminer ainsi les poules reproductrices susceptibles de transmettre les virus à leur descendance.

Outre la leucose lymphoïde, les ALVs induisent également des myélocytomatoses chez les poulets, ainsi qu'un spectre de tumeurs associées telles que des hémangiomes ou des néphroblastomes (Wu X, Qian K et al. 2010). Le virus responsable de cette affection a été identifié comme appartenant à un nouveau sous-groupe de virus leucosiques aviaires, l'ALV-J, qui résulte d'une recombinaison entre un virus exogène et un virus endogène (Payne LN 1998). L'ALV-J, contrairement aux ALVs des autres sous-groupes, présente la particularité d'être capable d'infecter toutes les lignées de poulets puisqu'aucune lignée de poulets résistante à l'infection virale n'a pu être identifiée (Venugopal K 1999).

ii) Les virus comportant des oncogènes ; parmi ces derniers il existe d'une part des virus réplicatifs, les virus du sarcome de Rous (RSV), et d'autre part des virus défectifs pour la multiplication virale :

- Le RSV induit la formation de tumeurs (sarcomes) chez le poulet quelques jours ou quelques semaines seulement après inoculation. En culture, le RSV peut induire la transformation de certaines cellules avec une grande efficacité, et les cellules transformées restent capables de produire des virus infectieux. Ce pouvoir de transformation provient du

gène *v-src* (*viral-sarc*) que le virus a acquis, selon toute vraisemblance, d'un rétrovirus proche de l'ALV qui aurait capturé et modifié accidentellement un *c-src* au cours de son cycle répliatif (Martin GS 2004).

- Les virus non répliatifs possèdent un ou deux oncogènes qui se sont insérés en place des séquences codantes. Ils se répliquent en présence de virus complet de type ALV tel le virus RAV.

Plusieurs familles de séquences endogènes apparentées aux ASLV ont été décrites dans le génome du poulet. On distingue: les ev (endogenous virus) dont les séquences sont proches de celles des virus exogènes, et des formes endogènes plus éloignées, les EAV (endogenous ancient virus), les ART-CH (avian retrotransposon from chicken genome) et les CR1 (chicken repeat 1) (Borisenko L 2003; Huda A, Polavarapu N et al. 2008).

Sous-groupe d'ALV	Virus prototype	Prévalence	Hôte	Exogène ou endogène	Récepteur (locus génétique chez l'hôte)	Pouvoir oncogène
A	RAV-1, RPL-12	Commun	Poulet	Exo	tv-a	Oui
B	RAV-2, MAV-2	Peu commun	Poulet	Exo	tv-b	Oui
C	RAV-49, RAV-7	Rare	Poulet	Exo	tv-c	Oui
D	RAV-50, CZAV	Rare	Poulet	Exo	tv-b	Oui
E	RAV-0, RAV-60	Commun	Poulet	Endo	tv-b/tv-e	Non
F	RAV-61		Faisan de colchide	Endo		
G	Virus du faisan doré		Faisan doré	Endo		
H	RAV-62		Perdrix hongroise	Endo		
I	Virus du colin de Gambel		Colin de Gambel	Endo		
J	HPRS-103, ADOL-Hc1	Commun	Poulet	Exo et endo	chNHE1	Oui

**Tableau 1. Classification et propriétés des différents sous-groupes d'ALV.** Adapté de Venugopal (Venugopal K 1999).

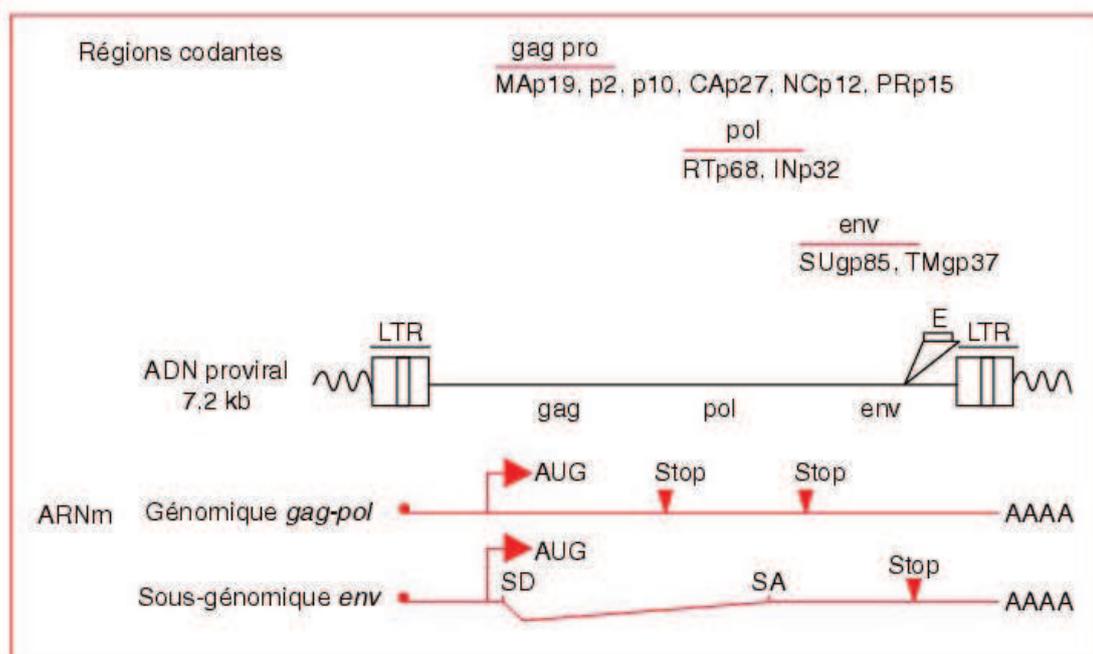
Les ALVs ont été classés en différents sous-groupes (A à J). Cette propriété repose sur la reconnaissance d'un récepteur spécifique par les différentes protéines d'enveloppe. Chez le poulet, on trouve les sous-groupes A, B, C, D, E et J (tableau 1), E et J étant constitués de virus endogènes. Les virus des sous-groupes F à I correspondant à des virus endogènes trouvés chez d'autres oiseaux que le poulet (faisan, perdrix).

### b) Caractéristiques morphologiques et génomiques des ALVs

Les ALVs présentent une morphologie caractéristique des rétrovirus du type C ; la localisation du virion est donc extracellulaire, et sa nucléocapside est ronde et centrale (figure 8). Quant au génome, il comporte les séquences classiques d'un génome simple (figure 9).



**Figure 8. Morphologie d'une particule ALV.** Vue au microscope électronique d'un virion en cours de bourgeonnement (A) et d'un virion mature (B) (de Harven E 1974). Les virions sont enveloppés et possèdent une forme sphérique. Leur taille est comprise entre 80 à 100 nm de diamètre.



**Figure 9. Organisation génomique d'ALV.** Adapté de Denesvre (Denesvre C and Dambrine G 2004). Schémas de l'ADN proviral et des ARN transcrits connus à partir de ce génome. (i) L'ADN proviral est constitué des LTRs qui contiennent les éléments de régulation de la transcription, ainsi que des régions codantes *gag*, *pol*, *env*. Les gènes *gag-pro* codent pour les protéines matrice (MAp19), p2, p10, capsid (Cap27), nucléocapside (NCp12) et protéase (PRp15). Le gène *pol* code pour la transcriptase inverse (RTp68) et l'intégrase (INp32). Le gène *env* code pour les glycoprotéines de surface (SUgp85) et transmembranaire (TMgp37) (ii) Les ARN transcrits représentés correspondent à l'ARN génomique et à l'ARN sous-génomique *env*. Certains éléments de ces ARN sont indiqués comme les sites donneur et accepteur d'épissage. SD : site donneur d'épissage, SA : site

accepteur d'épissage, AUG : codon d'initiation de la traduction, Stop : codon d'arrêt de la traduction, AAAA : queue polyadénylée.

## *B. Les rétrovirus endogènes porcins*

### *a) Généralités sur les PERVs*

Les virus endogènes porcins sont classés dans le genre gammarétrovirus. On distingue 3 classes de PERV (A, B et C) qui présentent des tropismes différents (Le Tissier P, Stoye JP et al. 1997; Akiyoshi DE, Denaro M et al. 1998).

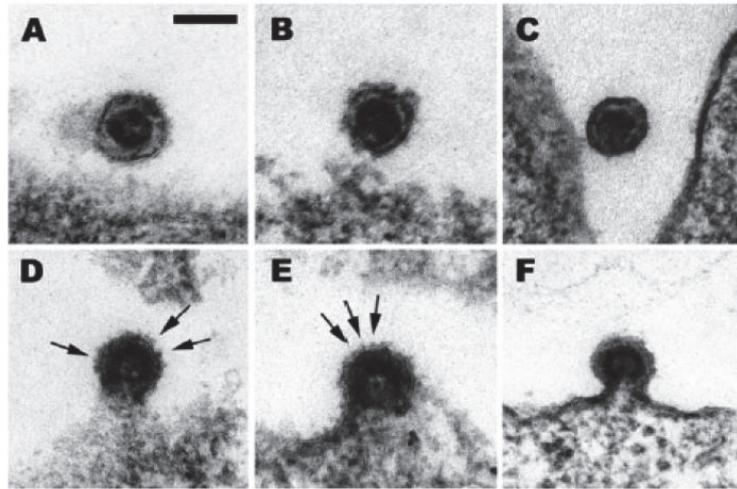
- PERV-A et -B sont présents dans les lignées germinales de tous les élevages testés jusqu'à présent (Le Tissier P, Stoye JP et al. 1997). Le nombre de copies de PERV-A est en général plus élevé que celui de PERV-B. On retrouve 30 à 50 copies de provirus dispersées dans le génome, dont la majorité est défective suite aux délétions et mutations accumulées. Les classes A et B sont polytropes, elles sont capables d'infecter, *in vitro*, des lignées cellulaires humaines et porcines (Takeuchi Y, Patience C et al. 1998; Patience C, Switzer WM et al. 2001).

- PERV-C, quant à lui, est présent en quantité moindre et dans certains cas absent des élevages (Patience C, Takeuchi Y et al. 1997; Bosch S, Arnauld C et al. 2000). La classe C est écotrope, le virus infecte uniquement les lignées cellulaires porcines.

- Il existe aussi un recombinant PERV-A/C généré par recombinaison entre un PERV-A et un PERV-C. Il est capable d'infecter des cellules humaines *in vitro* et se caractérise par un taux de réplication très élevé. Une autre caractéristique des PERVA/C est qu'aucune forme endogène de ce virus n'a été trouvée (Denner J 2008).

Notons que d'autres rétrovirus endogènes existent dans le génome porcine, et qui appartiennent au genre betarétrovirus. C'est le cas des séquences PMSN-1 et PMSN-4 présentes chez les porcs miniatures (Ericsson T, Oldmixon B et al. 2001).

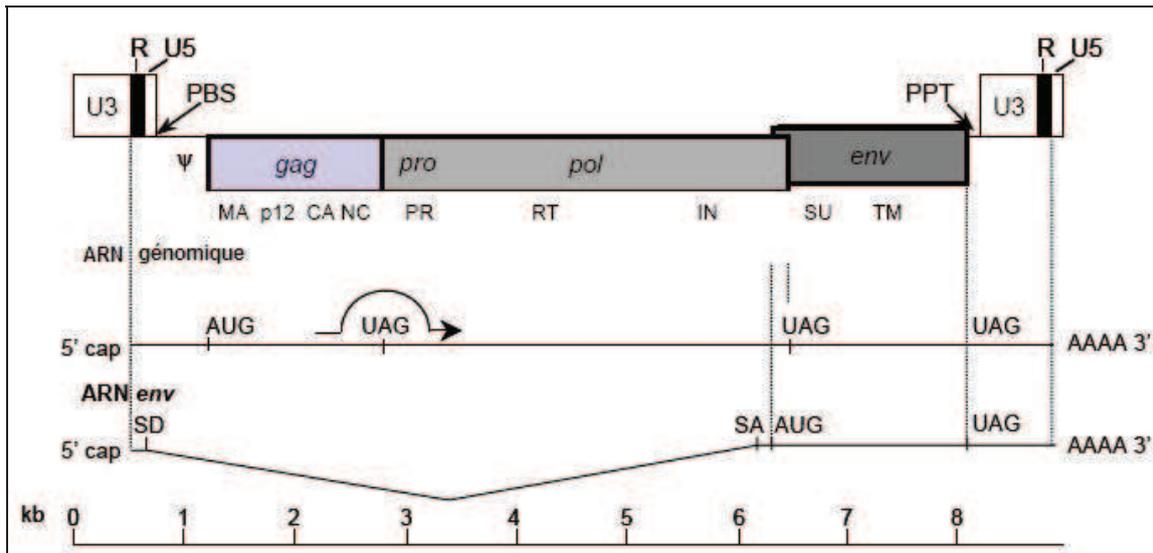
*b) Caractéristiques morphologiques et génomiques des PERVs*



**Figure 10. La morphologie des particules PERV.** D'après Preuss et al. (Preuss T, Fischer N et al. 2006), vues au microscope électronique des virions matures (A, B, et C) et en cours de bourgeonnement (D, E, et F). Les cadres A/D et B/E correspondent à des particules virales PERV-C issues de 2 lignées cellulaires porcines différentes, MAX-T et ST-IOWA respectivement. Les cadres C/F représentent des particules PERV-A/PERV-B provenant de la lignée cellulaire humaine HEK-293 infectée avec du surnageant de PK15. La barre dans le cadre A représente 100 nm.

Les virions PERV sont extracellulaires, ils présentent donc une morphologie caractéristique des rétrovirus du type C. Leur forme peut varier lors du bourgeonnement (figure 10) avec l'apparition transitoire de pics (indiqués par des flèches sur la figure 10) sur la surface de l'enveloppe pour les PERV-C alors que les PERV-A et -B en sont dépourvus (Preuss T, Fischer N et al. 2006).

Le génome des virus endogènes porcins possède les mêmes caractéristiques générales que les autres rétrovirus à génome simple (figure 11).



**Figure 11. Organisation génomique de PERV.** Adapté de Moalic (Moalic Y, Felix H et al. 2011). Schémas de l'ADN proviral et des ARN génomique et sous-génomique. Sur le schéma du provirus sont indiqués les LTR composés des fragments U3, R et U5, les régions codantes gag, pol, env, le signal d'encapsidation ψ et le PBS nécessaire à la transcription inverse.

Les ARN génomique (8,3 kb) et ARN env (3.1kb) résultent de la transcription de l'ADN proviral. Ils comportent une coiffe en 5' et une queue poly A en 3'. Certains éléments de ces ARN sont indiqués comme les sites donneur et accepteur d'épissage. SD : site donneur d'épissage, SA : site accepteur d'épissage, AUG : codon d'initiation de la traduction, Stop : codon d'arrêt de la traduction, AAAA : queue polyadénylée.

### c) Risques associés aux PERVs dans le cadre des xénotransplantations

Au cours de ces dernières années, de nombreux produits biologiques dérivés du porc (facteur VIII de la coagulation, pancréatine, héparine, etc.) ainsi que des organes et cellules pour la xénotransplantation, sont utilisés pour des applications cliniques chez l'homme (Yoo D and Giulivi A 2000; Schneider MK and Seebach JD 2011). Le développement de certains de ces produits a été considérablement freiné par la découverte, chez le porc, des rétrovirus endogènes dont un certain nombre correspondent à des provirus entiers.

Des questions se sont donc posées sur la possibilité d'infection des cellules humaines par PERV et sur les conséquences que puisse générer une telle infection. Ainsi, il a été démontré que certains PERVs sont capables d'infecter des cellules humaines *in vitro* (Patience C, Takeuchi Y et al. 1997). De plus, une recherche de PERV infectieux a été réalisée sur une lignée 293T sensible aux PERVs des sous-groupes A, B et C ; elle montre que seuls PERV-A et PERV-B se transmettent du plasma des mini-porcs NIH vers les cellules humaines, et que PERV-A est le seul capable de se répliquer et produire des virions après infection (Lee D, Kim NY et al. 2008).

Par ailleurs, une étude a été réalisée afin d'évaluer les effets d'une infection par PERV à long terme. Des cellules HEK-293 infectées ont été maintenues en culture durant 6 mois. Ceci a permis de mettre en évidence un faible taux de mutation dans les séquences LTRs et des différences insignifiantes dans la morphologie, croissance et apoptose des cellules après une infection à long terme (Yu P, Zhang P et al. 2009).

Dans certains cas, on a retrouvé de l'ADN de porc ou de PERV chez des receveurs humains, mais sans aucun signe de virémie ou d'infection (Yoo D and Giulivi A 2000; Switzer WM, Michler RE et al. 2001; Denner J 2011).

La xénotransplantation reste à nos jours la première cause d'inquiétude à propos de la transmission de PERV à l'homme. En cas d'infection de l'homme par PERV, cela pourrait entraîner des proliférations, des immunodéficiences ou des maladies neurodégénératives. Bien que le potentiel zoonotique de PERV demeure inconnu, le déséquilibre entre les demandes de greffes et les ressources disponibles est tel que le recours à la xénotransplantation reste toujours une solution envisageable.

## **Chapitre II : L'intégrase et l'intégration rétrovirale**

Dans ce chapitre, on présentera le processus d'intégration rétrovirale et l'intégrase. L'intégration représente une étape majeure du cycle rétroviral. Elle débute par la formation d'un complexe de préintégration (PIC), au sein duquel l'intégrase est transportée dans le noyau où elle réalise l'insertion du génome viral dans l'ADN de la cellule hôte.

### **I- Le transport nucléaire et le complexe de préintégration (PIC)**

Suite à l'entrée du corps viral dans le cytoplasme cellulaire, la capsid est rapidement désassemblée. L'ARN viral se trouve dans un complexe, dit complexe de rétrotranscription (RTC), qui contient certaines protéines virales. L'étape de transcription inverse de l'ARN viral en ADN double brin se fait de façon concomitante à la migration de ce complexe vers le noyau, et aboutit à la formation du PIC. L'ADNc se retrouve alors dans le PIC, constitué des protéines virales IN, MA, RT, NC et VPR (Bowerman B, Brown PO et al. 1989; Miller MD, Farnet CM et al. 1997) (cf. ci-dessous) et cellulaires (Suzuki Y and Craigie R 2002; Lin CW and Engelman A 2003; Llano M, Vanegas M et al. 2004). Parmi les protéines d'origine cellulaire qui co-purifient avec le PIC, on cite deux protéines largement étudiées ; BAF (Barrier to Autointegration) et HMGA1a (High Mobility Group protein 1a), qui interagissent avec l'ADN viral, et le facteur le plus connu LEDGF/p75, qui seront détaillés dans le chapitre 3.

Il était longtemps admis que, pour les oncorétrovirus, l'ADN viral n'accède au génome cellulaire qu'au moment de la mitose, lorsque l'enveloppe nucléaire est déstructurée, limitant leur capacité infectieuse à des cellules prolifératives. L'exemple le plus étudié concernait le passage au noyau du PIC de MLV dans des cellules mitotiques (Roe T, Reynolds TC et al. 1993; Yamashita M and Emerman M 2006). Ceci est remis en cause dans certains types cellulaires. En effet, il a été récemment démontré que le virus MLV et des vecteurs dérivés du MLV peuvent également infecter des neurones ainsi que des cellules neuroendocrines bloquées au niveau du cycle cellulaire (Liu XH, Xu W et al. 2011).

Dans le cas des ASLVs, l'import de l'ADNc dans les noyaux de cellules bloquées (cellules aviaires (DF-1), humaines (HeLa) et cellules de souris) au niveau du cycle cellulaire a été également démontré (Kukulj G, Katz RA et al. 1998; Katz RA, Greger JG et al. 2002).

Selon une étude récente, cet import ne dépendrait pas de la voie d'importine  $\alpha$  (Andrake MD, Sauter MM et al. 2008). D'autres auteurs ont cependant montré que l'IN d'ASV comporte une séquence NLS localisée dans le CCD (résidus 161-173) nécessaire pour l'import nucléaire (Levin A, Armon-Omer A et al. 2009).

Quant aux lentivirus, ils se distinguent des autres rétrovirus par un mécanisme d'import nucléaire du PIC qui leur permet d'infecter des cellules ne se divisant pas. Le PIC peut accéder donc au noyau de la cellule non proliférative en franchissant l'enveloppe nucléaire, à travers les pores nucléaires (CPN). Ces pores constituent un canal central traversant l'enveloppe nucléaire, et laissant diffuser les petites molécules dont la taille n'excède pas 40 à 60 kDa. Or, le PIC du VIH-1 possède une taille relativement importante, comprise entre 50 et 100 nm selon les auteurs. Le passage d'un complexe d'une telle taille se fait par un mécanisme actif mettant en jeu plusieurs éléments comme des séquences NLS et d'autres facteurs cités ci-dessous (pour une revue détaillée sur l'import nucléaire du PIC de VIH-1, voir (Rivière L, Darlix JL et al. 2010)).

Plusieurs facteurs peuvent faciliter le passage du PIC dans le noyau, comme la présence de la VPR capable d'induire des ruptures locales et transitoires dans l'enveloppe nucléaire (de Noronha CM, Sherman MP et al. 2001).

L'IN participe également à l'import nucléaire du PIC. Son association avec la molécule d'ADN viral est indispensable pour assurer l'import nucléaire de ce dernier (Ikeda T, Nishitsuji H et al. 2004). Il a été récemment montré que l'import nucléaire de l'IN de VIH-1 implique une interaction de son NLS localisée dans le cœur catalytique (résidus 161-173) avec l'importine  $\alpha$  (Levin A, Armon-Omer A et al. 2009). Le blocage de cette interaction affecte la réplication virale, permettant ainsi de confirmer le rôle de l'IN dans l'import nucléaire du PIC. Plus récemment, une interaction directe entre l'importine  $\alpha 3$  et l'IN de VIH-1 a été déterminée, et qui semble être requise pour l'import nucléaire de VIH-1 dans des cellules interphasiques ou non prolifératives (Ao Z, Danappa Jayappa K et al. 2010).

La matrice pourrait également participer à l'import nucléaire du PIC mais son implication reste très controversée. En effet, une séquence impliquée dans la localisation nucléaire de la matrice, via une interaction directe avec l'importine  $\alpha 1$ , a été trouvée dans la région N-terminale de la protéine. Cependant, l'implication de cette séquence dans l'import nucléaire du PIC a été ultérieurement controversée (Freed EO, Englund G et al. 1995; Reil H, Bukovsky AA et al. 1998; Mannioui A, Nelson E et al. 2005; Hearps AC, Wagstaff KM et al.

2008 ). Plus tard, une autre séquence NLS a été identifiée dans la région C-terminale de la matrice (Haffar OK, Popov S et al. 2000). Les résultats rapportés par ces auteurs indiquent que la mutation des deux séquences N- et C-terminales de matrice se traduit par un défaut de réplication du VIH-1 dans des macrophages.

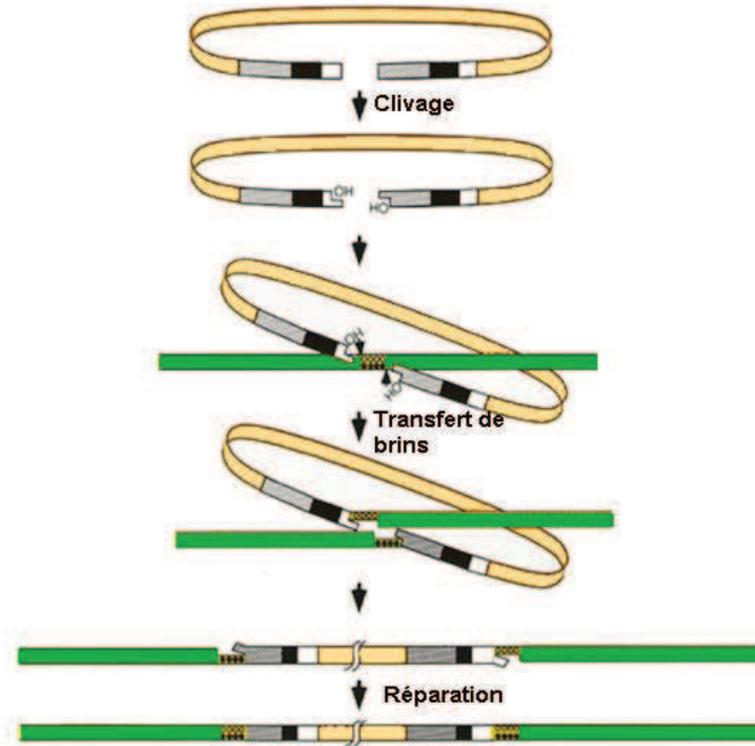
Par ailleurs, une séquence triple brin de 99 nucléotides a été identifiée au centre de l'ADN du HIV, appelée *DNA flap central*, et dont l'absence cause une accumulation de l'ADN viral au voisinage de la membrane nucléaire (Zennou V, Petit C et al. 2000). Notons que certaines études indiquent que le *DNA flap central* n'est pas essentiel pour la réplication du virus (Marsden MD and Zack JA 2007). D'autre part, certains auteurs suggèrent que la séquence *DNA flap* est impliquée dans la décapsidation de VIH-1, qui aurait lieu au niveau de la membrane nucléaire (Arhel NJ, Souquere-Besse S et al. 2007).

Enfin, une implication de la séquence cPPT-CTS, identifiée auparavant comme un signal impliqué dans la transcription inverse, a été récemment signalée dans l'import nucléaire de VIH-1 (Rivière L, Darlix JL et al. 2010).

## **II- Le mécanisme d'intégration**

La réaction d'intégration est catalysée par l'intégrase. Cette enzyme se fixe aux extrémités LTR de l'ADN viral, mature les extrémités par clivage d'un dinucléotide terminal en 3' (3'-processing), et intègre de façon covalente l'ADN viral dans le génome hôte. Le mécanisme d'intégration se termine par une troisième et dernière étape de réparation des lésions chromosomiques; cette étape est réalisée par des protéines cellulaires impliquées dans la réparation de l'ADN. Il en résulte une petite duplication des séquences cellulaires de part et d'autres du virus, dont la taille est spécifique du virus (6 pb pour ASLV, 5 pb pour VIH-1).

La réaction complète d'intégration est brièvement schématisée sur la figure 12.



**Figure 12. Mécanisme d'intégration**, adapté de Freed (Freed EO 2001). Sont représentées les étapes de clivage, de transfert de brins et de réparation. Le clivage correspond à l'élimination de 2 nucléotides adjacents aux CA terminaux par l'intégrase. Le transfert de brins consiste en une liaison entre les extrémités 3' de l'ADN viral et 5' de l'ADN cellulaire. Ce processus se termine par une étape de réparation, réalisée par les protéines cellulaires, qui permet de combler les brèches provoquées par l'insertion de l'ADN viral.

#### *A. Clivage en 3' et transfert de brins*

L'ADN rétroviral se présente avec des extrémités franches correspondant aux séquences LTRs (Guntaka RV, Mahy BW et al. 1975; Varmus HE, Guntaka RV et al. 1975). Chaque extrémité virale comporte une séquence CANN dont le dinucléotide CA est très conservé entre les différents rétrovirus.

L'IN se fixe spécifiquement sur les extrémités de l'ADN viral et, par élimination du dinucléotide NN, libère une extrémité 5' CA-OH. Cette réaction de clivage est appelée «3'-processing» (Engelman A, Mizuuchi K et al. 1991; Vink C, Yeheskiely E et al. 1991). En termes de spécificité, une mutation de la séquence CA abolit totalement l'activité de l'IN tandis que l'intégrité des séquences adjacentes est de moindre importance (Esposito D and Craigie R 1998). Cette réaction de clivage a lieu dans le cytoplasme des cellules infectées immédiatement après l'étape de transcription inverse; elle nécessite la présence de molécules d'eau (Engelman A, Mizuuchi K et al. 1991).

Notons que certaines intégrases rétrovirales sont également capables de couper une séquence palindromique interne correspondant à une jonction LTR-LTR que l'on retrouve dans les formes circulaires de l'ADN viral (Delelis O, Parissi V et al. 2007).

Le transfert de brins conduit ensuite à l'insertion covalente de l'ADN viral dans le génome de la cellule infectée. Cette seconde réaction a lieu de façon concomitante pour les deux extrémités de l'ADN viral avec un décalage précis de paires de bases entre les deux points d'insertion opposés. Elle correspond à une trans-estérification en une seule étape qui met en jeu une attaque nucléophile réalisée par le groupement CA-OH de l'ADN viral sur une liaison phosphodiester de l'ADN cellulaire, ce qui signifie qu'il n'existe pas d'intermédiaire covalent entre l'enzyme et l'ADN (Engelman A, Mizuuchi K et al. 1991; Vink C, Yeheskiely E et al. 1991).

*In vitro*, il est possible de reproduire le processus d'intégration d'une seule extrémité en utilisant des oligonucléotides double-brin (ODNs) marqués, d'environ 15-20 paires de bases, ou plus comme dans le cas de MLV (Jonsson CB, Donzella GA et al. 1993), mimant la séquence des extrémités virales des LTR U5 ou U3 en présence de l'enzyme IN recombinante. On peut visualiser l'activité de coupure endonucléotique ou de transfert de brins par électrophorèse (Delelis O, Carayon K et al. 2008). Les deux réactions, 3'-processing et transfert de brin, requièrent *in vitro* un cofacteur métallique ( $Mn^{2+}$ ,  $Mg^{2+}$  ou  $Zn^{2+}$ ) (Engelman A and Craigie R 1995; Agapkina J, Smolov M et al. 2006; Carayon K, Leh H et al. 2010). Or, c'est certainement le  $Mg^{2+}$  qui intervient *in vivo*.

### *B. Désintégration*

Enfin, l'inverse de la réaction de transfert de brin peut également être observé *in vitro* (Chow SA, Vincent KA et al. 1992). Cette réaction est appelée la désintégration ; elle peut être catalysée par l'IN entière ou par le seul CCD. Aucune donnée expérimentale n'indique aujourd'hui que cette réaction puisse se produire *in vivo*.

### *C. Réparation*

Une dernière réaction, qui consiste en une réparation de l'ADN au site d'intégration, complète le mécanisme d'intégration. Il en résulte une duplication d'une courte séquence chromosomique de part et d'autre de la séquence virale. Bien qu'une étude ait suggéré qu'une

activité ADN polymérase-ADN dépendante de l'intégrase puisse être impliquée dans cette étape (Acel A, Udashkin BE et al. 1998), les facteurs DNA-PK, PARP-1 et NBS1 seraient responsables de la réparation selon certains auteurs (Brin E, Yi J et al. 2000), mais ceci n'est pas confirmé (Ariumi Y, Turelli P et al. 2005).

- La DNA-PK (DNA-dependent protein kinase) est un composant clé du mécanisme de NHEJ (Non-homologous end joining). DNA-PK est un complexe multiprotéique constitué d'une sous-unité catalytique DNA-PKcs et d'un hétérodimère Ku70/Ku86 qui reconnaît et fixe les extrémités d'ADN double brin (Lees-Miller SP and Meek K 2003). L'implication de ce facteur dans l'étape de réparation a été proposée suite à une étude de l'intégration rétrovirale dans des cellules dont le gène codant la protéine DNA-PK a été muté. L'infection des cellules mutantes défectives en DNA-PK s'accompagne d'un taux élevé d'apoptose. Le signal apoptotique a été attribué à la présence d'intermédiaires d'intégration non réparés (Skalka AM and Katz RA 2005). Notons que le rôle de la DNA-PK dans la réplication virale ne s'arrête pas à l'étape de réparation ; en effet, selon une étude récente, la DNA-PK est impliquée dans l'expression des gènes viraux de HIV-1, dû à son interaction directe avec l'ARN pol II (Tyagi S, Ochem A et al. 2011).

- Poly(ADP-ribose) polymerase 1 est aussi une protéine impliquée dans la réparation de l'ADN. Après la fixation à l'ADN contenant un « gap », PARP-1 utilise NAD<sup>+</sup> pour catalyser l'attachement de poly(ADP-ribose) à une variété d'enzymes de réparation d'ADN. Puis, elle se dissocie de l'ADN laissant l'accès libre à d'autres protéines de réparation. L'implication de PARP-1 dans l'étape d'intégration rétrovirale a été montrée lors d'une étude menée par Gaken et ses collègues, qui montre qu'une inhibition de l'activité de PARP bloque l'infection des cellules de HeLa, NIH 3T3 (« mouse fibroblasts ») et L1210 (« mouse lymphoblastoid cells ») par les vecteurs rétroviraux (Gaken J, Tavassoli AM et al. 1996). D'autres auteurs ont suggéré que PARP-1 est requise pour l'intégration du génome de HIV-1 au niveau des centromères, dans des cellules humaines et murines (Kameoka M, Nukuzuma S et al. 2005). A part son implication dans l'intégration, un autre rôle a été suggéré pour PARP-1 dans la réplication virale, notamment au cours de la transcription des gènes viraux. En effet, Parent et ses collègues ont montré que PARP-1 entre en compétition avec Tat ou avec le complexe Tat/p-TEFb (positive transcription elongation factor b) au niveau de l'ARN TAR, et agit ainsi comme un régulateur négatif de la transcription de VIH-1 (Parent M, Yung TM et al. 2005).

- NBS1 (Nijmegen breakage syndrome-1 protein) est également une protéine qui agit au niveau des cassures d'ADN (Lukas C, Falck J et al. 2003), au sein d'un complexe contenant la kinase MRE11 (meiotic recombination 11 homologue), et RAD50 (Tauchi H, Kobayashi J et al. 2001; Kobayashi J, Tauchi H et al. 2002). De plus, NBS1 peut également recruter la kinase ATM (ataxia telangiectasia mutated) aux sites de cassure (Falck J, Coates J et al. 2005) ; or, l'implication de cette dernière dans la réparation de l'ADN n'est toujours pas claire (Daniel R, Katz RA et al. 2001; Ariumi Y, Turelli P et al. 2005; Daniel R and Pomerantz RJ 2005; Dehart JL, Andersen JL et al. 2005; Lau A, Swinbank KM et al. 2005).

Récemment, une étude a montré que NBS1 se lie à l'ADN viral de VIH-1 et recrute l'ATM kinase, induisant ainsi une intégration stable de l'ADN viral et empêchant une apoptose due à la présence d'intermédiaires d'intégration non réparés (Smith JA, Wang FX et al. 2008).

#### *D. L'intégration concertée*

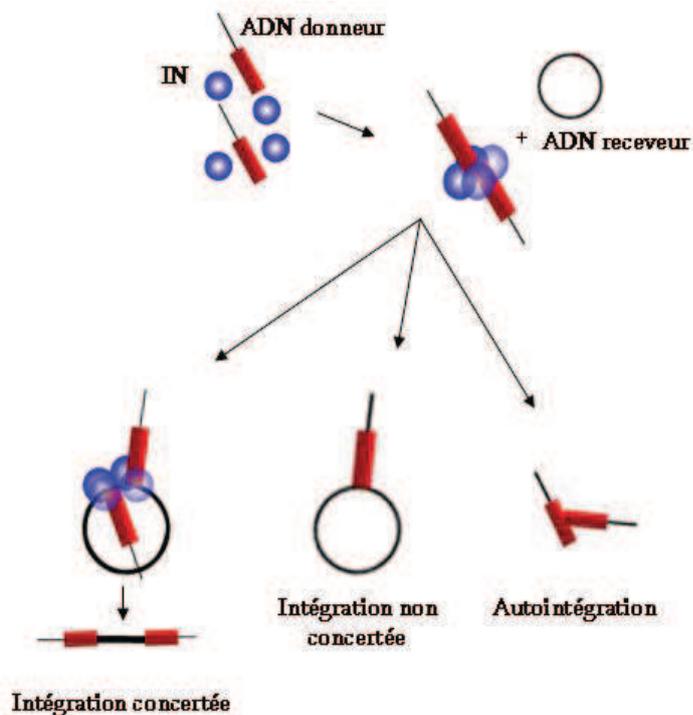
L'intégration correspond à l'insertion simultanée des deux extrémités d'un ADN viral sur un même site de l'ADN cellulaire ; on parle donc d'intégration concertée. La reproduction de cette réaction *in vitro* nécessite l'incubation d'un ADN donneur composé d'un gène de sélection (*supF*) et, au minimum, des 10-20 nucléotides terminaux des séquences LTR (séquences *att*), avec un ADN receveur circulaire, représentant l'ADN cible, en présence de l'IN recombinante purifiée (Goodarzi G, Im GJ et al. 1995; Hindmarsh P and Leis J 1999; Yang F and Roth MJ 2001; Moreau K, Faure C et al. 2003; Sinha S and Grandgenett DP 2005).

L'intégration concertée est difficile à reproduire efficacement *in vitro*, et la réaction peut donner naissance, à part les produits d'intégration concertée, à deux autres produits: (i) des produits d'intégration non concertée correspondant à l'insertion d'une ou de plusieurs extrémités provenant d'un ou de plusieurs ADN donneur, sur différents sites de l'ADN receveur, et (ii) des produits d'auto-intégration correspondant à l'insertion d'un ADN donneur dans un autre (figure 13).

Les études ont montré qu'il existe plusieurs facteurs capables d'augmenter l'efficacité de la réaction. Des protéines virales, tel la NCp7 (Carteau S, Batson SC et al. 1997; Carteau S, Gorelick RJ et al. 1999; Gao K, Gorelick RJ et al. 2003; Cruceanu M, Urbaneja MA et al. 2006) ou des partenaires protéiques d'origine cellulaire, peuvent affecter l'efficacité de la réaction d'intégration concertée, comme les protéines cellulaires HMG I(Y) ou BAF (Harris

D and Engelman A 2000). L'exemple le plus frappant est représenté par la protéine LEDGF/p75 qui stimule fortement l'intégration réalisée par l'IN du VIH-1.

Dans une réaction d'intégration *in vitro*, LEDGF/p75 favorise l'intégration d'une seule molécule d'ADN viral dans l'ADN cible, quand la réaction est réalisée sans un ordre précis dans l'addition des protéines et des substrats. Au contraire, une réaction d'intégration concertée efficace des 2 extrémités d'un ADN viral est observée quand les différents composants de la réaction sont ajoutés dans l'ordre suivant : le complexe IN-ADN viral préformé, suivi d'une addition de LEDGF/p75 (Pandey KK, Sinha S et al. 2007; Raghavendra NK and Engelman A 2007). LEDGF/p75 stabilise donc la liaison de l'IN à l'ADN, stimulant ainsi l'intégration concertée. D'autres études sont venues plus tard confirmer le rôle de LEDGF/p75 dans l'intégration concertée (cf. § Oligomérisation de l'IN, et § Structure cristallographique de l'IN).



**Figure 13. Modes d'intégration.**

(i) intégration concertée correspondant à l'insertion simultanée des deux extrémités de l'ADN donneur au niveau du même site de l'ADN receveur. (ii) intégration non concertée résultant de l'insertion d'une extrémité d'un ADN donneur dans l'ADN receveur. (iii) auto-intégration qui résultant de l'insertion d'un ADN donneur dans un autre.

### III- L'intégrase

#### A. Généralités sur l'intégrase

L'intégration de l'ADN rétroviral dans le génome de la cellule hôte est catalysée par l'enzyme virale IN.

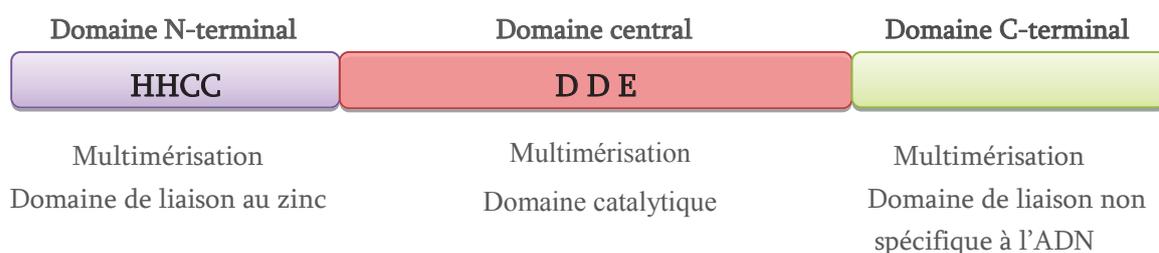
### a) Rôles de l'IN dans le cycle de réplication virale

En plus des activités catalytiques citées ci-dessus (cf. § Le mécanisme d'intégration), l'IN possède des propriétés non catalytiques essentielles pour le cycle de réplication.

Des études ont montré que certains mutants de l'IN, présentant des activités catalytiques identiques à la protéine sauvage, affectent d'autres étapes de la réplication virale. Ainsi, plusieurs mutations de l'IN bloquent la réplication virale au niveau de l'étape de transcription inverse (Lu R, Limon A et al. 2004; Ao Z, Fowke KR et al. 2005; Tsurutani N 2000). Ce phénotype pourrait être expliqué par un changement dans la structure du complexe de rétrotranscription. L'IN peut également agir indirectement sur la transcription inverse en recrutant des protéines cellulaires, comme la SIP1/Gemin2 (détaillée dans le chapitre suivant) (Nishitsuji H, Hayashi T et al. 2009).

Un autre rôle de l'IN concerne son implication dans l'import nucléaire du PIC (cf. § Le transport nucléaire et le complexe de préintégration (PIC)). Bien que son mécanisme d'action dans cette étape du cycle viral reste controversé.

Enfin, l'intégrase peut être indirectement impliquée dans la régulation de la transcription de l'ADN proviral. Suite au processus d'intégration, l'IN se trouve liée à l'ADN intégré, et sera dégradée par la voie proteasome-ubiquitine. Elle interagit donc avec la VBP1 (von Hippel-Lindau binding protein 1), un partenaire du complexe Cul2/VHL impliqué dans le processus de polyubiquitylation, qui semble être également impliqué dans la régulation de la transcription des gènes viraux (Mousnier A, Kubat N et al. 2007).



**Figure 14. Organisation des domaines de l'intégrase.** L'IN est constituée de 3 domaines: (i) le domaine N-terminal fixe les ions zinc via les résidus HHCC (ii) le domaine central contient le site catalytique D,D,E, et reconnaît l'ADN viral (iii) le domaine C-terminal responsable de la fixation non spécifique de l'ADN. Ces 3 domaines jouent également un rôle dans la multimérisation de la protéine.

## *b) Structure en domaines de l'IN*

Des études de digestion par des protéases, et des analyses de comparaison de séquences de plusieurs INs ont permis de démontrer que l'IN est composée de trois domaines fonctionnels distincts (Johnson MS, McClure MA et al. 1986; Andrade MD and Skalka AM 1996) (figure 14).

### - Le domaine N-terminal (NTD)

Il est constitué approximativement des 50 premiers acides aminés. Le NTD contient un motif HHCC (H : histidine, C : cystéine) conservé dans les INs de tous les rétrovirus et rétrotransposons eucaryotiques (Khan E, Mack JP et al. 1991; Andrade MD and Skalka AM 1996); La taille de ce domaine est plus importante chez les gammarétrovirus, notamment chez MLV, où 50 acides aminés précèdent le domaine HHCC (Johnson MS, McClure MA et al. 1986).

Ce domaine est impliqué dans la fixation des ions zinc. Cependant, la longueur de l'hélice entre les histidines et les cystéines est plus longue que celle dans les doigts de zinc de la majorité des facteurs de transcription. La structure du NTD de l'IN de VIH-1 a été résolue par RMN (spectroscopie à résonance magnétique nucléaire) (Cai M, Zheng R et al. 1997) et par cristallographie, et a montré que la fixation des ions zinc permet de stabiliser la configuration des hélices alpha de ce domaine, et favorise la multimérisation de la protéine (Zheng R, Jenkins TM et al. 1996). Concernant le rôle du NTD, diverses études ont été faites montrant que le rôle majeur de ce domaine est son implication dans la multimérisation de l'IN. De plus, ce domaine semble jouer un rôle indirect dans la fixation de l'ADN viral (Carayon K, Leh H et al. 2010).

### - Le domaine core catalytique (CCD)

Il constitue une région centrale d'approximativement 150 acides aminés, caractérisé par un motif DDE (D : acide aspartique, E : acide glutamique), sachant qu'entre le second acide aspartique et l'acide glutamique, il existe 35 acides aminés. Ce domaine est essentiel pour l'activité catalytique de l'IN (Drelich M, Wilhelm R et al. 1992; Engelman A and Craigie R 1992). La résolution de la structure du CCD de l'IN de VIH-1 a permis de mettre en évidence son existence en état dimérique et sa capacité à fixer des ions  $Mg^{2+}$ , nécessaires au

fonctionnement de l'enzyme (Dyda F, Hickman AB et al. 1994). Ce domaine a été ensuite cristallisé en complexe avec des ions  $Mg^{2+}$ , confirmant ainsi les résultats précédents (Goldgur Y, Dyda F et al. 1998). Le CCD est aussi impliqué dans la multimérisation de l'IN. Dans le cas de ASLV, la mutation du résidu N197C favorise l'intégration non concertée *in vitro* suite à une perte de la forme dimérique de l'IN (Moreau K, Faure C et al. 2004). Une autre fonction décrite pour ce domaine est la reconnaissance de l'ADN viral (Jenkins TM 1997) et l'ADN cellulaire (Appa RS, Shin CG et al. 2001).

#### - Le domaine C-terminal (CTD)

C'est le moins conservé des trois. Ce domaine est constitué d'approximativement 70 à 100 acides aminés. Trois rôles ont été décrits pour le CTD. Au début, des analyses de mutation et de délétion des INs de HIV-1 et ASV montre que le CTD possède une activité de fixation à l'ADN (Lutzke RA, Vink C et al. 1994). Ensuite, la structure d'un fragment isolé du CTD de l'IN de HIV-1 a été obtenue par RMN (Lodi PJ, Ernst JA et al. 1995) et révèle la présence d'un domaine SH3 présent sous forme d'un dimère en solution (Eijkelenboom AP, Lutzke RA et al. 1995). Ce domaine SH3 pourrait être impliqué dans la fixation de peptides ou de protéines. Enfin, des études biochimiques ont montré la présence dans ce domaine de résidus impliqués dans la multimérisation de l'IN de ASV et de VIH-1. Dans le cas de VIH-1, une suppression du domaine CTD réduit fortement la capacité de formation de structures oligomériques de la protéine (Lutzke RA and Plasterk RH 1998). Dans le cas d'ASLV, Moreau et al. ont démontré que la mutation du résidu L240A réduit la dimérisation de l'IN et sa capacité à catalyser une intégration concertée (Moreau K, Faure C et al. 2003). Pour MLV, ce domaine contient une insertion de 36 acides aminés supplémentaires (Johnson MS, McClure MA et al. 1986) ; la mutation de plus de 28 acides aminés du CTD supprime l'activité catalytique de l'IN (Jonsson CB, Donzella G A et al. 1996).

#### *B. Oligomérisation de l'IN*

En solution, l'IN peut exister sous forme monomérique, dimérique, tétramérique ou même dans des états oligomériques d'ordre supérieur. De nombreux facteurs peuvent influencer l'état d'oligomérisation de la protéine en solution : la concentration en substrats d'ADN, en sel, en protéines, les interactions qu'elle peut avoir avec des cations, la présence de détergents ou de glycérol (Deprez E, Tauc P et al. 2000; Leh H, Brodin P et al. 2000; Deprez E, Tauc P et al. 2001; Villanueva RA, Jonsson CB et al. 2003). *In vivo*, l'intégrase

fonctionne sous forme tétramérique afin de pouvoir effectuer l'intégration concertée (Bera S, Pandey KK et al. 2009).

D'abord, il a été démontré que certaines combinaisons de mutants individuellement inactifs de l'intégrase, peuvent avoir une activité *in vitro* (Engelman A, Bushman FD et al. 1993; van Gent DC, Vink C et al. 1993; van den Ent FM, Vos A et al. 1999). Il a été donc suggéré que l'IN agit sous forme multimérique. Cette hypothèse a été ensuite renforcée par les résultats de différentes études qui ont suggéré qu'un dimère d'IN est requis à chacune des deux extrémités de l'ADN viral durant la réaction de clivage et que ce dimère peut réaliser le transfert de brins d'une seule extrémité du LTR (Bao KK, Wang H et al. 2003; Guiot E, Carayon K et al. 2006). De plus, il a été démontré qu'une intégration concertée efficace nécessite une forme multimérique de l'IN (Sinha S, Pursley MH et al. 2002), et suite à des tests d'oligomérisation *in vitro*, Faure et al. ont montré, que seule la forme tétramérique de l'IN est capable de réaliser l'intégration concertée (Faure A, Calmels C et al. 2005). Plus tard, des données obtenues sur la structure de l'IN en complexe avec l'ADN viral sont venues supporter le fait qu'un tétramère d'intégrase se fixe sur l'ADN viral (Baranova S, Tuzikov FV et al. 2007; Ren G, Gao K et al. 2007; Hare S, Gupta SS et al. 2010).

Certains facteurs peuvent contribuer à la stabilisation de ce complexe, comme la protéine cellulaire LEDGF/p75. Une étude utilisant la méthode « mass spectrometry protein footprinting » a montré, qu'il y a un changement dans la conformation de l'IN de HIV au niveau du segment flexible entre le CCD et CTD, suite à l'addition de l'ADN viral (Zhao Z, McKee CJ et al. 2008). Au contraire, un tel changement dans la conformation de l'IN n'est pas observé en présence de LEDGF/p75 (McKee CJ, Kessl JJ et al. 2008). Ceci suggère que LEDGF/p75 module différemment les structures de l'IN libre ou de l'IN en complexe avec l'ADN viral. Ces observations ont été confirmées par deux études structurales de l'intasome menées par Hare et ses collègues, et Michel et ses collègues (Hare S, Di Nunzio F et al. 2009; Michel F, Crucifix C et al. 2009) (cf. § Structure cristallographique de l'IN). puis par une étude très récente utilisant la technique de FRET (Förster Resonance Energy Transfer) dont les résultats montrent deux conformations différentes du tétramère d'IN en présence de l'ADN viral ou de LEDGF/p75 (Kessl JJ, Ignatov M et al. 2011).

Enfin, une déstabilisation de l'oligomérisation de l'IN peut empêcher son assemblage avec l'ADN. Ainsi, la mutation de résidus impliqués dans la multimérisation de l'IN affecte fortement l'efficacité de la réaction d'intégration (Moreau K, Faure C et al. 2003; Moreau K, Faure C et al. 2004). Une autre étude a montré qu'un peptide dérivé de LEDGF/p75 peut se

lier spécifiquement à l'IN de VIH-1 et modifier son état d'oligomérisation, d'un dimère actif à un tétramère inactif, incapable de catalyser la réaction de clivage en 3' (Hayouka Z, Rosenbluh J et al. 2007).

### *C. Structure cristallographique de l'IN*

Les premières structures cristallographiques obtenues des INs ne comportaient que des domaines isolés ou deux domaines sur les trois CCD/CTD d'une part et NTD/CCD d'autre part.

Ces différentes structures montrent une conservation des structures monomériques entre virus et une conservation de l'association entre les CCDs (association par les hélices  $\alpha 1$  et 5) (Bujacz G, Jaskólski M et al. 1995; Lodi PJ, Ernst JA et al. 1995).

Récemment, une association différente du dimère du CCD a été obtenue au sein de notre équipe lyonnaise qui montre une association à trois hélices au lieu de deux (Ballandras A, Moreau K et al. 2011).

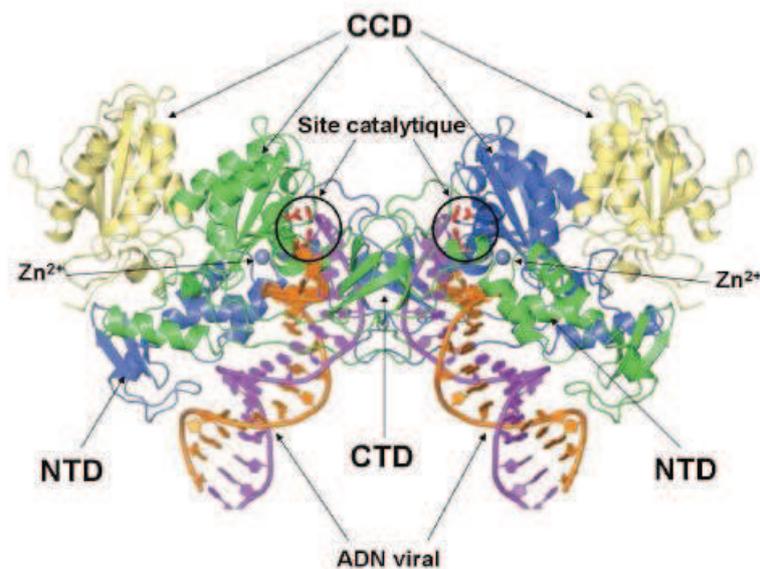
La disposition des deux CTD est différente dans les structures des INs de VIH-1, RSV et SIV (Eijkelenboom AP, Sprangers R et al. 1999; Chen Z, Yan Y et al. 2000; Yang ZN, Mueser TC et al. 2000).

Pour ce qui concerne l'association du domaine NTD au CCD, si un dimère du domaine NTD est toujours présent au sein de la structure du double-domaine CCD/NTD, les surfaces d'interaction sont nettement différentes par rapport au dimère du domaine isolé (Dyda F, Hickman AB et al. 1994). Ceci suggère une flexibilité des domaines CCD et NTD par rapport au domaine CCD.

Récemment, des modèles 3D de l'IN de VIH-1 seule, ou en complexe avec LEDGF/p75 en présence d'ADN viral ou cellulaire, ont été obtenus par microscopie électronique (Michel F, Crucifix C et al. 2009). Les résultats obtenus suggèrent que l'unité fonctionnelle est constituée de deux dimères d'intégrases asymétriques en complexe avec deux molécules de LEDGF/p75. Dans ce modèle, le NTD ne dimérise pas, mais se trouve en contact avec le CTD de la même molécule d'IN. En présence de l'ADN, il y a un changement conformationnel du NTD et du CTD. Ainsi, une seule molécule IN de chaque dimère est impliquée dans la liaison à l'ADN viral, et procède ainsi aux réactions de clivage et transfert de brins, tandis que l'autre molécule n'est impliquée que dans le positionnement de l'ADN viral au sein du site actif du dimère opposé.

Une structure en 3D de l'IN entière de PFV (Prototype Foamy virus) présente au sein d'un complexe avec l'ADN viral (figure 15) (Hare S, Gupta SS et al. 2010 ), et l'ADN cellulaire (Maertens GN, Hare S et al. 2010) a été également résolue, permettant de confirmer qu'un tétramère d'IN est nécessaire pour réaliser une intégration concertée. Ces travaux montrent surtout une association inattendue du tétramère, dans laquelle l'interaction entre les CCDs telle que observée jusque là n'a plus la même importance. Les résultats de la première étude montrent également qu'il existe une forte interaction entre l'IN et l'ADN viral, capable de résister à une incubation prolongée dans des conditions de forces ioniques élevées. La structure 3D entre l'IN entière de PFV et l'ADN viral révèle une sous-unité interne contenant un dimère d'intégrase en liaison stable avec une molécule d'ADN viral, et une sous-unité externe constituée d'une paire d'intégrase symétriquement reliée à la sous-unité interne, formant ainsi un tétramère. L'interface du dimère est stabilisée par des interactions NTD-CCD (sous-unité intérieure, sous-unité extérieure). Par contre, la structure obtenue du complexe IN PFV-ADN cellulaire révèle un espace entre les dimères d'IN au sein de l'intasome qui facilite la fixation de l'ADN cellulaire en multiple conformations. En se basant sur la structure de l'intasome de PFV comme modèle, une modélisation de la structure de l'intasome de VIH-1 a été réalisée (Krishnan L, Li X et al. 2010), et qui a révélé l'importance du rôle inconnu jusqu'à présent de l'Arg20 et Lys266 dans la fixation de l'ADN. Pour une revue détaillée récente sur les structures en 3D de l'IN, voir (Li X, Krishnan L et al. 2011).

Par ailleurs, un modèle de l'architecture d'un dimère de l'IN entière d'ASV a été déterminé récemment par SAXS (small angle X-ray scattering) et cross-link. A l'inverse des structures résolues par cristallographie des différents domaines de l'IN, où l'interaction entre les CCD semble jouer un rôle important, dans ce modèle, un dimère d'IN entière n'est stable que par interaction CTD-CTD et par les interactions entre le NTD d'une sous-unité avec le CCD et le CTD de l'autre sous unité du dimère (Bojja RS, Andrade MD et al. 2011).



**Figure 15. Structure en 3D de l'IN du Foamy Virus en association avec l'ADN viral (Hare S, Gupta SS et al. 2010 ).** Cette structure représente un tétramère d'intégrase, dont les monomères en bleu et vert se trouvent à l'intérieur du complexe, directement liés à l'ADN viral, alors que les monomères en jaunes sont à l'extérieur du complexe. Les sites catalytiques impliqués dans liaison à l'ADN viral sont également indiqués. CCD : central core domain, CTD : C-terminal domain, NTD : N-terminal domain.

#### *D. Les inhibiteurs de l'IN*

La protéine intégrase constitue une cible idéale pour la découverte de thérapies antirétrovirales. Les données obtenues sur ses activités catalytiques, ont permis le développement d'inhibiteurs par deux approches : soit par identification (i) d'inhibiteurs de la réaction « 3'-processing » ; ces derniers ciblent l'IN empêchant sa fixation à l'ADN viral, et sont appelés INBI (IN DNA-Binding Inhibitors) (Deprez E, Barbe S et al. 2004), ou par identification (ii) d'inhibiteurs de la réaction de transfert de brins, INSTI (IN Strand Transfer Inhibitors), qui ciblent le complexe IN-ADN viral préformé et empêchent sa liaison à l'ADN cellulaire (Hazuda DJ, Felock P et al. 2000).

Cependant, il n'y a que les INSTI qui ont montré un potentiel antiviral. C'est essentiellement le cas de Raltegravir (Isentress®), qui a été développé en se basant sur une étude menée par Hazuda et al. (Hazuda DJ, Felock P et al. 2000), et qui a été approuvé pour l'utilisation clinique en automne 2007 comme la première drogue antirétrovirale dirigée contre l'intégrase. Des études réalisées *in vitro* et *in vivo* ont montré que ce composé a une activité sur des souches virales résistantes aux différents antirétroviraux utilisés jusqu'à

présent : les inhibiteurs nucléosidiques, nucléotidiques et non nucléosidiques de la transcriptase inverse, les inhibiteurs de la protéase et les inhibiteurs de fusion. Cependant, comme avec les autres classes d'antirétroviraux, la résistance à cet inhibiteur a rapidement été mise en évidence. Cette résistance est due à la sélection et à l'émergence, sous anti intégrase, de variants viraux initialement minoritaires portant des mutations de résistance sur le gène de l'intégrase et se multipliant sous la pression de sélection de cet inhibiteur. Il a été donc conseillé d'associer le raltegravir à d'autres antirétroviraux actifs afin d'éviter l'apparition rapide de variants viraux résistants à cette nouvelle classe (Nguyen BY, Isaacs RD et al. 2011).

À l'heure actuelle, l'interaction entre LEDGF/p75 et l'intégrase constitue une nouvelle cible séduisante et prometteuse pour la conception de nouveaux inhibiteurs antirétroviraux (Du L, Zhao Y et al. 2008 ; Fan X, Zhang FH et al. 2011).

## **IV- Les sites d'intégration**

### *A. Les profils d'intégration*

Il est actuellement établi que les rétrovirus sont partagés en 3 groupes selon leurs préférences d'intégration (Mitchell RS, Beitzel BF et al. 2004).

Le premier groupe est constitué des rétrovirus dont les profils d'intégration montrent une préférence pour les gènes actifs. On distingue : le VIH-1; en effet, une analyse des sites préférentiels d'intégration du VIH-1 dans des lignées cellulaires humaines montre que 69% de l'ensemble des intégrations ont lieu dans les gènes transcriptionnellement actifs (Schröder AR, Shinn P et al. 2002; Lewinski MK, Bisgrove D et al. 2005), le SIV (Crise B, Li Y et al. 2005), le VIH-2 et le FIV (Kang Y, Moressi CJ et al. 2006).

Le second groupe contient les rétrovirus (MLV, PERV et FV) qui favorisent l'intégration dans les sites d'initiation de la transcription et au niveau des îlots CpG. En effet, environ 20% des événements d'intégration de MLV sont observés au niveau des extrémités 5' des unités de transcription (Wu X, Li Y et al. 2003). De plus, environ 17% des intégrations ont lieu dans une fenêtre de 2 kb autour des îlots CpG, alors que ces fenêtres ne représentent que 2% du génome. Une fréquence élevée d'intégration (11%) a été également signalée au niveau des sites 'DNase I-hypersensitive' (Lewinski MK, Yamashita M et al. 2006). Pareil pour PERV, les résultats ont montré une forte similitude avec MLV, permettant ainsi de suggérer que les gammarétrovirus possèdent un profil commun concernant le choix des sites d'intégration (Moalic Y, Blanchard Y et al. 2006; Moalic Y, Félix H et al. 2009). De même, une intégration au niveau des sites d'initiation de la transcription, et une préférence pour les îlots CpG ont été observées pour le rétrovirus FV (foamy virus) (Trobridge GD, Miller DG et al. 2006).

Le dernier groupe inclut les ASLVs, HTLV-1 et MMTV, et présentent une intégration aléatoire. Ainsi pour les ASLVs, il a été montré que le niveau d'intégration dans les gènes est faible (40%), et que ces rétrovirus ne favorisent pas les sites d'initiation de la transcription (Mitchell RS, Beitzel BF et al. 2004; Narezkina A, Taganov KD et al. 2004). De plus, un niveau élevé de transcription semble inhiber l'intégration de l'ASV dans les gènes (Weidhaas JB, Angelichio EL et al. 2000; Maxfield LF, Fraize CD et al. 2005). Les rétrovirus aviaires présentent donc un profil d'intégration plutôt aléatoire ; Idem pour les rétrovirus HTLV-1 et MMTV (Derse D, Crise B et al. 2007; Faschinger A, Rouault F et al. 2008).

## B. Les paramètres de sélection

Le choix du site d'intégration de l'ADN viral au sein du génome cellulaire dépend de différents paramètres (Lewinski MK, Yamashita M et al. 2006; Ciuffi A 2008):

- Structure de la chromatine : des études *in vitro*, mettant en jeu l'IN purifiée avec de l'ADN nu comme ADN cible, ont montré que certaines DNA-BP (DNA-binding protein) sont capables en se liant à l'ADN d'entraver l'accès à l'IN, empêchant ainsi l'intégration (Pryciak PM and Varmus HE 1992). Au contraire, le remplacement de l'ADN nu par de l'ADN associé à des nucléosomes favorisait l'intégration. La distorsion ou l'enroulement de l'ADN semblaient induire une courbure favorable à l'intégration (Pryciak PM, Sil A et al. 1992; Pryciak PM and Varmus HE 1992; Pruss D, Bushman FD et al. 1994). Ces études révélaient donc une certaine préférence dans la sélection des sites d'intégration, liée à la structure de l'ADN. Cependant, comme l'ADN cible existe sous forme de chromatine, ces résultats ne représentaient pas les faits réels qui pourraient avoir lieu dans une cellule infectée. Ainsi l'utilisation d'un ADN cible associé à 13 nucléosomes qui pourrait être compacté en chromatine par addition de l'histone H1 a permis de montrer que la structure de la chromatine affecte la sélection des sites d'intégration du VIH-1 et de l'ASV de façons différentes. La compaction de l'ADN semble stimuler l'intégration par l'IN de l'ASV, et défavoriser l'intégration par l'IN de VIH-1 (Taganov KD, Cuesta I et al. 2004). De même Lesbats et ses collègues ont montré que la compaction du polynucléosome réduit l'accessibilité de l'ADN à l'IN inhibant l'intégration concertée réalisée par l'IN du VIH-1 (Lesbats P, Botbol Y et al. 2011). Ainsi, une infection non productive pourrait donc être le produit d'une intégration dans des régions d'hétérochromatine (Jordan A, Defechereux P et al. 2001; Jordan A, Bisgrove D et al. 2003). Or, il a été montré que pour VIH-1, l'intégration dans les régions d'hétérochromatines est un phénomène rare *in vivo* et elle n'est responsable qu'en partie de la latence (Han Y, Lassen K et al. 2004; Lewinski MK, Bisgrove D et al. 2005). De plus, Shan L et ses collègues ont montré récemment que l'interférence transcriptionnelle peut être un des principaux facteurs responsables de la maintenance de la latence du VIH-1, en montrant une préférence d'intégration dans la même orientation transcriptionnelle du gène hôte (Shan L, Yang HC et al. 2011). Enfin, l'étude menée par Lesbats et ses collègues a montré également que, bien que certaines structures de la chromatine puissent être réfractaires pour l'intégration, le remodelage du nucléosome par le complexe SWI/SNF, mettant en jeu une interaction

directe entre l'INI1 (integrase interactor 1) et l'IN de VIH-1, permet de franchir cette barrière naturelle.

- La transcription : La transcription peut également jouer un rôle dans la sélection des sites d'intégration : (i) quand un ADN est transcriptionnellement actif, l'intégration peut être bloquée à cause d'un encombrement stérique dû à la présence du complexe d'ARN polymérase II, (ii) l'intégrase requiert un ADN double brin pour exercer l'intégration, ce qui n'est pas le cas durant une transcription active, (iii) la transcription peut induire un remodelage de la chromatine en une structure moins favorable pour l'intégration. Ce qui explique que les centromères sont défavorables pour l'intégration, alors que les îlots CpG ou les séquences transcrites sont favorables. Ainsi, il a été montré que la présence de sites d'intégration dans des régions promotrices, par exemple au niveau des éléments MAR (matrix attachment regions) peut faciliter l'intégration et la transcription rétrovirale (Kulkarni A, Pavithra L et al. 2004). La préférence d'intégration dans les séquences transcrites a été montrée par Maxfield et ses collègues, qui ont comparé la fréquence et la distribution de l'intégration d'un rétrovirus aviaire dans le gène métallothionéine avant, et après l'induction par ZnSO<sub>4</sub> (Maxfield LF, Fraize CD et al. 2005).

- L'intégrase : l'IN est considérée comme un déterminant principal du profil d'intégration (Lewinski MK, Yamashita M et al. 2006). La construction d'intégrases chimériques entre le VIH et le FIV (Shibagaki Y and Chow SA 1997) a permis de révéler l'implication du domaine CCD et plus particulièrement de l'acide aminé 119 de son hélice  $\alpha 2$  (Appa RS, Shin CG et al. 2001) dans le choix des sites d'intégration (Harper AL, Sudol M et al. 2003). De la même manière, un acide aminé en position analogue chez l'IN d'ASV a été identifié, et qui affecte le choix des sites d'intégration (Harper AL, Sudol M et al. 2003).

D'autre part, l'IN peut interagir avec des protéines cellulaires capables d'influencer la sélection des sites d'intégration (cf. chapitre 3, Les cofacteurs cellulaires de l'intégration), comme le facteur de transcription LEDGF/p75, qui interagit avec l'IN de VIH-1 et favorise le ciblage des gènes actifs (Ciuffi A, Llano M et al. 2005) (cf. § LEDGF/p75, chapitre 3). Notons que pour les INs de PERV et d'ASLV (nos modèles d'étude), aucun cofacteur cellulaire de l'IN impliqué dans la sélection des sites d'intégration n'a été, jusqu'à présent, identifié.

### *C. La sélection des sites d'intégration et la thérapie génique*

Les rétrovirus, et particulièrement les VIH et MLV, sont utilisés en tant que vecteurs dans les thérapies géniques (Fischer A, Hacein-Bey-Abina S et al. 2011). Or, l'intégration des vecteurs viraux dans certaines régions du chromosome est parfois accompagnée durant ces thérapies de l'apparition de leucémies dans le modèle animal et chez l'homme (Hacein-Bey-Abina S, Von Kalle C et al. 2003; Alexander BL, Ali RR et al. 2007; Bushman FD 2007). Afin d'éviter l'intégration dans certains sites, l'utilisation de vecteurs non intégratifs est alors envisagée (Wanisch K and Yáñez-Muñoz RJ 2009). Une autre stratégie utilisée consiste à cibler l'intégration dans des sites spécifiques du génome ; ainsi, l'intégrase fusionnée à un doigt de zinc polydactil favorise 7 à 10 fois l'intégration sur un site spécifique (Giri MS, Nebozhyn M et al. 2006), de même la surexpression de protéines chimères contenant le CTD de l'IN fusionnée au NTD de la protéine HPA $\alpha$  modifie les sites d'intégration des vecteurs HIV (Silvers RM, Smith JA et al. 2010). Les progrès récents dans la caractérisation des profils d'intégration des rétrovirus vont donc permettre de prévenir les effets adverses liés à l'utilisation de ces vecteurs et permettre d'élargir l'application thérapeutique de cette puissante stratégie (Bushman FD 2007; Qasim W, Gaspar HB et al. 2009; Biffi A, Bartolomae CC et al. 2011).

## **Chapitre III : Les cofacteurs cellulaires de l'intégration**

Bien que l'intégrase recombinante seule soit nécessaire et suffisante pour réaliser l'intégration *in vitro*, de nombreuses études indiquent que le virus recrute des cofacteurs cellulaires lors de l'intégration *in vivo* ; certains agissent sur l'ADN viral au sein du PIC (cf. § I- Les constituants cellulaires du PIC), d'autres interagissent directement avec l'intégrase (cf. § II- Les partenaires cellulaires de l'IN de VIH-1). La plupart des partenaires cellulaires de l'IN, dont le rôle dans la réplication virale a été élucidé, sont des cofacteurs de l'IN du VIH-1. Récemment les cofacteurs cellulaires de l'IN de MLV ont été identifiés par double hybride, mais dont le rôle dans la réplication virale reste inconnu (cf. § III- Les partenaires cellulaires de l'IN de MLV). Parmi ceux-ci, le facteur Brd2 auquel on va s'intéresser dans cette étude. L'identification de tels cofacteurs cellulaires et la compréhension de leurs rôles dans l'intégration faciliteront le développement de thérapies antivirales basées sur l'inhibition de l'intégration.

En ce qui concerne les cofacteurs cellulaires des INs d'ASLV et de PERV, aucune recherche de partenaires de l'IN de PERV n'a été réalisée jusqu'aujourd'hui. Par contre, pour ASLV, une seule protéine cellulaire, Daxx (death associated protein), a été identifiée comme partenaire de cette IN (Greger JG, Katz RA et al. 2005) (cf. § Daxx).

### **I- Les constituants cellulaires du PIC**

#### *A. HMGA1a*

HMGA1a (High Mobility Group chromosomal protein A1) a été isolée à partir du PIC de VIH-1 (Farnet CM and Bushman FD 1997). Les membres de cette famille possèdent la capacité de lier l'ADN et les protéines. Concernant son rôle dans l'intégration, HMGA1a semble stimuler l'intégration concertée de VIH-1 et ASV, *in vitro* (Hindmarsh P, Ridky T et al. 1999). Elle pourrait se lier aux LTRs, rapprochant ainsi les deux extrémités virales et induisant la décondensation des LTRs, ce qui en faciliterait l'accès à l'IN. Par contre, son rôle dans l'intégration *in vivo* est remis en question (Beitzel B and Bushman F 2003). Une étude récente a démontré que HMGA1a serait impliquée dans la régulation de l'épissage spécifique des sites de l'ARN de VIH-1 (Tsuruno C, Ohe K et al. 2011).

## *B. BAF*

En utilisant des PICs isolés de cellules infectées par MLV ou VIH-1 dans une réaction d'intégration *in vitro*, on observe une préférence d'intégration de l'ADN viral dans l'ADN cible ajouté à la réaction, plutôt que dans l'ADN viral lui-même. Ainsi, un facteur cellulaire, barrière d'autointégration (BAF), qui empêche cette autointégration suicidaire a été isolé du PIC de MLV (Lee MS and Craigie R 1994), puis du PIC de VIH-1 (Lin CW and Engelman A 2003).

BAF interagit avec l'ADN (Zheng R, Ghirlando R et al. 2000), interagit directement avec les histones H3 et H1.1 (Montes de Oca R, Lee KK et al. 2005), mais également avec des protéines nucléaires possédant un domaine LEM conservé (LAP2, Emerin, MAN1), en présence ou en absence d'ADN (Lee KK, Haraguchi T et al. 2001; Shumaker DK, Lee KK et al. 2001; Suzuki Y, Yang H et al. 2004). Ces résultats suggèrent que le complexe chromatine/BAF/emerin/MAN1 joue un rôle crucial dans l'assemblage nucléaire post-mitotique et dans la décondensation de la chromatine. Le fait que BAF puisse compacter l'ADN (Skoko D, Li M et al. 2009) pourrait expliquer l'inaccessibilité de cet ADN pour l'autointégration.

## **II- Les partenaires cellulaires de l'IN de VIH-1**

Un nombre croissant de partenaires de l'IN du VIH-1 a été identifié ces dernières années. Tous ne jouent pas un rôle au cours du processus intégratif lui-même, mais peuvent intervenir au cours d'autres étapes du cycle viral. On distingue ainsi, ceux qui sont impliqués dans différentes étapes pré-intégratives, ceux qui sont impliqués dans l'intégration proprement parlé, et ceux qui sont impliqués dans les étapes post-intégratives. Parmi ces facteurs, LEDGF/p75 (cf. ci-dessous) constitue un facteur majeur et a fait l'objet de nombreux travaux ces dernières années.

### *A. Les cofacteurs impliqués dans les étapes pré-intégratives*

#### *a) Gemin2*

Gemin2 est un membre du complexe SMN (Survival Motor Neuron), qui est impliqué dans l'assemblage des snRNPs (Spliceosomal small Nuclear Ribonucleoproteins). Gemin2 a été identifiée par double hybride comme partenaire cellulaire de l'IN de HIV-1 (Hamamoto S, Nishitsuji H et al. 2006). Cette interaction implique une région de 101 acides aminés (résidus

137-238) de la protéine Gemin2 et le domaine C-terminal de l'intégrase. Ce facteur stimule *in vitro* l'assemblage de l'IN et la RT au niveau de l'ARN génomique afin d'activer la transcription inverse (Nishitsuji H, Hayashi T et al. 2009).

#### *b) Transportin/SR2*

La protéine Transportin-SR2 (TRN-SR2 ou TNPO3) est un membre de la famille importin- $\beta$ , qui reconnaît les régions riches en sérine/arginine dans les facteurs d'épissage des pré-ARNm, et transporte ensuite ces facteurs vers le noyau.

La Transportin-SR2 a été identifiée comme un cofacteur de la réplication de VIH-1 par deux dépistages ARNi du génome (« genome-wide RNAi screens ») (Brass AL, Dykxhoorn DM et al. 2008; König R, Zhou Y et al. 2008), et l'a été confirmé par une étude récente (Krishnan L, Matreyek KA et al. 2010). Une interaction directe entre TRN-SR2 et l'IN de VIH-1 a été montrée par double hybride et validée par pull down (Christ F, Thys W et al. 2008). Cette protéine serait impliquée dans l'import nucléaire de VIH-1, comme le montre des expériences de KO du gène. En effet, le KO de TRN-SR2 (i) réduit la formation des ADN circulaires (ADN formés dans le noyau des cellules) (ii) réduit le pourcentage des particules virales observées au niveau de la membrane nucléaire (Albanese A, Arosio D et al. 2008).

Cependant, des travaux récents suggèrent que le rôle de la TRN-SR2 dans l'infection par VIH-1 impliquerait plutôt la CA, bien qu'à ce jour, aucune interaction n'ait été démontrée entre la capsid virale et la TRN-SR2 (Krishnan L, Matreyek KA et al. 2010).

#### *c) Importin*

Imp7 est l'une des nombreuses importines qui se lient aux protéines ribosomiques des cellules des mammifères, et jouent un rôle dans l'import nucléaire. Des essais d'import nucléaire réalisés dans un système reconstitué *in vitro* ont permis de montrer que l'Imp7 contribue à l'import du complexe de reverse transcription de VIH-1 vers le noyau (Fassati A, Görlich D et al. 2003). Par des essais de co-immunoprécipitation, une interaction entre l'imp7 et l'IN de VIH-1 a été montrée, et par des analyses de mutagenèse, deux régions dans le domaine C-terminal de l'IN importantes pour la liaison avec l'imp7, ont été déterminées (Ao Z, Huang G et al. 2007).

Récemment, une autre importine (l'Imp $\alpha$ 3) a été identifiée comme un partenaire de l'IN de VIH-1. Un knockdown (KD) par shRNA (short hairpin RNA) de l'imp $\alpha$ 3 diminue

significativement l'infection de VIH-1 dans des cellules HeLa, CD4(+) C8166T et des macrophages primaires. Des analyses par RT-PCR ont montré que le KD de l'Imp $\alpha$ 3 n'a aucun effet sur la transcription inverse du virus mais réduit le taux d'ADN viral circulaire suggérant un rôle de l'Imp $\alpha$ 3 dans l'import nucléaire de VIH-1 (Ao Z, Danappa Jayappa K et al. 2010). Deux régions dans le CTD de l'IN de VIH-1 sont requises pour son interaction avec l'Imp $\alpha$ 3 et dans l'import nucléaire de l'ADN viral (Jayappa KD, Ao Z et al. 2011).

## *B. Les cofacteurs impliqués dans l'intégration*

### *a) SNF5/INI1*

Le facteur INI1 (Intégrase Interactor 1) est l'homologue humain du facteur de transcription SNF5 chez la levure. SNF5/INI1 est une sous-unité du complexe SWI/SNF. Ce complexe, très conservé chez les eucaryotes, est impliqué dans le remodelage des structures de la chromatine réfractaires pour l'intégration (Lesbats P, Botbol Y et al. 2011).

SNF5/INI1 est le premier co-facteur cellulaire de l'IN de VIH-1 à avoir été identifié, par double hybride. Son interaction avec l'IN faciliterait la liaison de cette dernière avec l'ADN (Kalpana GV, Marmon S et al. 1994).

D'autres données suggèrent que SNF5/INI-1 interviendrait à différentes étapes du cycle de la réplication virale :

- Une première série d'expériences suggère que SNF5/INI1 pourrait participer à une réponse cellulaire antivirale. En effet, peu après son entrée dans la cellule, le virus induit la relocalisation des protéines SNF5/INI1 et PML du noyau vers le cytoplasme, où elles vont co-localiser avec le PIC pour être ensuite réimportées dans le noyau (Turelli P, Doucas V et al. 2001).

- L'expression de l'INI-1 dans des cellules INI-1<sup>null</sup> (cellules MON) provoque une augmentation de la production des particules virales (Yung E, Sorin M et al. 2001); le fragment S6, le domaine minimal permettant l'interaction de INI-1 avec l'IN, est impliqué dans les étapes tardives de la réplication virale, essentiellement l'assemblage et le bourgeonnement du virus (Yung E, Sorin M et al. 2001).

- Des résultats obtenus par PCR quantitative montrent une accumulation des formes circulaires à 2 LTR suite à un KO de SNF5/INI-1 (Maroun M, Delelis O et al. 2006), suggérant un contrôle de INI-1 dans le contrôle de la recombinaison des formes circulaires.

- Il a été également démontré que INI1 peut également jouer un rôle dans les étapes post-intégratives en participant à la régulation de la transcription de VIH-1 par le biais de la protéine Tat (Ariumi Y, Serhan F et al. 2006; Tréand C, du Chéné I et al. 2006).

Enfin, le rôle de l'INI1 semble être multifactoriel dans le cycle et reste encore assez peu clair.

#### *b) P300*

P300 est une acétyltransférase cellulaire qui module la conformation de la chromatine par acétylation des histones. L'interaction entre p300 et l'IN de VIH-1 a été montrée *in vitro* et dans les cellules, par la technique de FRET (Fluorescence Resonance Energy Transfer Technique) (Cereseto A, Manganaro L et al. 2005 ). Ces auteurs ont suggéré que p300 serait impliquée dans l'acétylation de l'IN, et que cette acétylation faciliterait l'activité de « transfert de brins » de l'IN en augmentant son affinité pour l'ADN. Les mutations des sites d'acétylation de l'IN répriment la réplication du virus en inhibant le transfert de brins (Cereseto A, Manganaro L et al. 2005 ). Une seconde étude a confirmé l'implication de p300 dans l'acétylation de l'IN, mais contredit le fait que cette acétylation soit indispensable pour la réplication virale (Topper M, Luo Y et al. 2007). Très récemment, il a été démontré que la reconnaissance de l'acétylation de l'intégrase par la protéine KAP1, membre de la famille Trim, induit une déacétylation en formant un complexe contenant la déacétylase HDAC1, et inhibe l'intégration (Allouch A, Di Primio C et al. 2011). Ainsi, le débat sur l'importance de l'acétylation de l'intégrase dans la réplication virale continue et nécessite des analyses complémentaires.

#### *c) LEDGF/p75*

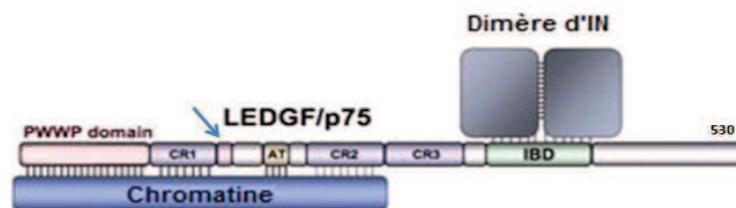
LEDGF/p75 «Lens Epithelial Derived Growth Factor/p75 » est une protéine nucléaire, intimement liée aux chromosomes. LEDGF est un membre de la famille des Hepatoma-derived growth factor related proteins (HRP).

L'interaction entre LEDGF/p75 et l'IN de HIV-1 a été découverte séparément par deux groupes, en analysant des protéines associées à l'IN de VIH-1 dans des cellules humaines (Cherepanov P, Maertens G et al. 2003; Turlure F, Devroe E et al. 2004), et en recherchant les interactants de l'IN par double hybride (Emiliani S 2005). L'interaction entre les deux partenaires met en jeu un petit domaine d'environ 80 résidus qui se trouve dans la

région C-terminale de LEDGF/p75, appelé IBD (IN-Binding Domain) (Cherepanov P, Devroe E et al. 2004) (figure16) et la région couvrant les résidus 165-173 de l'IN d'VIH-1. La mutation du résidu Q168 n'affecte pas l'activité enzymatique de l'IN mais uniquement son interaction avec LEDGF/p75 (Emiliani S 2005; Busschots K 2007).

Actuellement, LEDGF/p75 est reconnue comme un partenaire spécifique des intégrases des lentivirus, et un des constituants cellulaires du PIC (Llano M, Vanegas M et al. 2004; Cherepanov P 2007).

Récemment, deux études structurales de l'intasome menées par Hare et ses collègues, et Michel et ses collègues (Hare S, Di Nunzio F et al. 2009; Michel F, Crucifix C et al. 2009) ont montré un changement de la conformation de l'IN en présence de LEDGF/p75. De plus, une étude plus récente utilisant la technique de FRET (Förster Resonance Energy Transfer) dont les résultats montrent deux conformations différentes du tétramère d'IN en présence de l'ADN viral ou de LEDGF/p75 (Kessl JJ, Ignatov M et al. 2011).



**Figure 16. Organisation des domaines de LEDGF/p75 (Ciuffi A, Llano M et al. 2005).** LEDGF/p75 contient 530 acides aminés. On distingue plusieurs domaines fonctionnels : (i) Un domaine de 92 résidus en N-ter contenant un motif très conservé PWWP (Pro-Trp-Trp-Pro) qui agit comme un domaine d'interaction protéine-protéine (Stec I, Nagl SB et al. 2000), et comme un domaine de fixation à l'ADN (DBD) (Qiu C, Sawada K et al. 2002). (ii) Deux copies d'un motif appelé crochet AT, qui se lient aux régions de l'ADN riches en AT, ont été retrouvées entre les résidus 178-183 et 191-197. La capacité de LEDGF/p75 à s'associer à la chromatine nécessite la coopération du domaine PWWP et du crochet AT (Turlure F, Maertens G et al. 2006). (iii) Un domaine de fixation à l'IN (IBD) constitué de 80 acides aminés et qui se trouve dans la région C-terminale (Cherepanov P, Devroe E et al. 2004). De plus, une séquence de localisation nucléaire (NLS) a été trouvée entre les résidus 148-156 (indiquée par une flèche).

Un nombre important d'études a souligné l'importance de LEDGF/p75 dans la réplication de HIV-1 ;

- La mutation des résidus de l'IN impliqués dans son interaction avec LEDGF/p75 (R166 et Q168) bloque la réplication virale au niveau de l'étape d'intégration (Emiliani S 2005).

- *In vitro*, l'addition de LEDGF/p75 dans une réaction d'intégration, stimule l'intégration en favorisant l'affinité de l'IN pour l'ADN (Cherepanov P 2007).

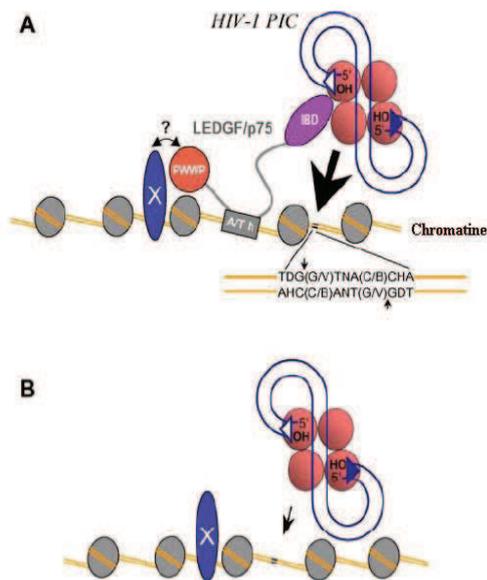
- La réduction du niveau d'expression de LEDGF/p75 par RNAi génère trois conséquences très importantes: (i) dans les cellules en interphase, la majorité de l'IN se trouve localisée dans le cytoplasme des cellules plutôt que dans le noyau en présence de LEDGF/p75, (ii) les INs se dissocient du chromosome, (iii) cette redistribution est accompagnée d'une réduction du taux des INs. Ces observations montrent que LEDGF/p75 est un cofacteur important de l'IN, impliqué dans la stabilité de l'IN (protection contre l'ubiquitination et la dégradation par le protéasome), dans la localisation de l'IN au noyau et dans l'association de l'IN à la chromatine. En signalant que l'IN est aussi, la cible d'une dégradation par le système ubiquitine-protéasome cellulaire (Shun MC, Raghavendra NK et al. 2007).

- La déplétion partielle de cellules humaines en LEDGF/p75 conduit à une sélection des sites d'intégration du VIH-1 différente par rapport aux cellules normales. L'intégration du VIH-1 se faisant normalement dans les unités de transcription, celle-ci devient plus fréquente dans les régions riches en GC en l'absence de LEDGF/p75 (Ciuffi A, Llano M et al. 2005). Par ailleurs, l'analyse des sites d'intégration de VIH-1 dans des cellules de souris KO pour le gène codant LEDGF/p75 a confirmé sa capacité à diriger l'intégration dans les unités transcriptionnelles, loin des régions promotrices et des îlots CpG (Shun MC, Raghavendra NK et al. 2007; Ciuffi A 2008).

- Inversement, des expériences de sur-expression de LEDGF/p75 ont été également réalisées; Une surexpression de la région IBD de LEDGF/p75 fusionnée avec la GFP (Green Fluorescent Protein) entraîne une inhibition de la réplication virale dans des cellules HeLaP4 et MT-4, au niveau de l'intégration (De Rijck J, Vandekerckhove L et al. 2006).

La surexpression d'une forme dominante négative de LEDGF/p75 a permis la sélection de virus résistants dont l'intégrase a été mutée (résidus A128 et E170). Bien que l'affinité de l'IN mutée pour LEDGF/p75 a été réduite, le virus résistant est toujours dépendant de LEDGF/p75 pour sa réplication, confirmant ainsi le rôle primordial de ce facteur cellulaire (Hombrouck A, De Rijck J et al. 2007).

Selon le modèle admis actuellement (Engelman A and Cherepanov P 2008), LEDGF/p75 associée à la chromatine agirait comme un récepteur pour le PIC et dirigerait l'intégration vers le locus génomique le plus proche (figure 17). En absence de LEDGF/p75, les PICs restent capables d'engager une voie indépendante du facteur LEDGF/p75.



**Figure 17. Modèle représentant le Rôle de LEDGF/p75 dans l'intégration de HIV-1 (Engelman A and Cherepanov P 2008).** Le nucléosome, l'unité de base de la chromatine, est constitué par des histones représentés en ovales de couleur gris, associés aux chromosomes représentés par des lignes oranges. (A) LEDGF/p75 interagit d'une part avec l'IN par son domaine IBD, et d'autre part avec l'ADN cellulaire via son domaine PWWP (capable de reconnaître un facteur X non identifié encore) et ses crochets AT. L'interaction avec LEDGF/p75 dirige ainsi l'intégration du PIC du VIH-1 dans des sites préférentiels des chromosomes. L'intégration peut se faire au sein d'une séquence palindromique préférentielle décrite pour VIH-1 (dont la séquence est indiquée). (B) En absence de LEDGF/p75, les PICs restent capables de s'intégrer mais de façon indépendante de l'intégration dirigée par LEDGF/p75 et du facteur X.

#### d) YY1

YY1 (YinYang 1) est un co-facteur cellulaire du PIC. Son interaction avec les INs de VIH-1, MLV et ASV, a été montrée par co-immunoprécipitation et pull down. L'immunoprécipitation de la fraction cytoplasmique de cellules infectées par MLV, suivie d'un southern blot et de chIP (chromatin immunoprecipitation) ont montré que YY1 s'associe avec l'ADN viral de MLV. *In vitro*, YY1 stimule les activités catalytiques des trois INs. Une déplétion du gène dans les cellules NIH 3T3 réduit l'intégration de MLV *in vivo* (Inayoshi Y, Okino Y et al. 2010).

### C. Les cofacteurs impliqués dans les étapes post-intégratives

#### a) VBPI

VBPI (Von Hippel Lindau binding protein 1), ou préfoldine3, constitue une sous-unité d'un complexe constitué de 6 préfoldines. La préfoldine se lie à des protéines cibles non natives, comme l'actine ou la tubuline, et les transfère à la chaperonine cytosolique qui facilite alors leur repliement (Vainberg IE, Lewis SA et al. 1998). Elle pourrait également

s'associer à la protéine pVHL permettant ainsi la polyubiquitination des protéines cibles et leur dégradation par le protéasome 26S (Tsuchiya H, Iseda T et al. 1996).

La protéine VBP1 est un cofacteur cellulaire de l'IN identifié par criblage double hybride (Rain JC, Cribier A et al. 2009). Elle jouerait un rôle entre les étapes d'intégration et de transcription. En effet, une inhibition de l'expression de VBP1 diminue la polyubiquitylation de l'intégrase et provoque sa stabilisation. Ainsi, l'interaction de la VBP1 avec l'IN permettrait la dégradation, par le système protéasome-ubiquitine, de l'IN qui se trouverait toujours liée à l'ADN viral après son intégration. La fixation de l'IN par VBP1 va également entraîner le recrutement du complexe Cul2/VHL (cullin2-based von Hippel-Lindau ubiquitin ligase), dont VBP1 fait partie. Ce complexe pourrait être impliqué dans la régulation de la transcription des gènes viraux, suggérant ainsi un rôle indirect de l'IN dans la phase post-intégrative (Mousnier A, Kubat N et al. 2007).

#### *b) EED*

EED, produit du gène *eed* (Embryonic Ectoderm Development gene), agit comme un répresseur de la transcription.

EED a été initialement identifié comme étant un partenaire de la protéine MA de VIH-1 par criblage double hybride (Peytavi R, Hong SS et al. 1999), puis comme un partenaire de la protéine Nef (Witte V, Laffert B et al. 2004; Rakotobe D, Tardy JC et al. 2007).

Des analyses de mutagenèse et des essais pull down ont mis en évidence des sites dans le domaine C-terminal de la protéine EED (résidus 212-264) capables de fixer l'IN (Violot S, Hong SS et al. 2003). De plus, en examinant la distribution cellulaire de l'IN par IEM (ImmunoElectron Microscopy), EED co-localise avec l'IN et la MA (Violot S, Hong SS et al. 2003). Notons qu'une étude structurale récente a permis de caractériser l'interaction IN-MA-EED (Rakotobe D, Violot S et al. 2008).

EED pourrait stimuler l'intégration *in vitro*, et la co-localisation EED-MA-IN permet de suggérer un rôle de EED dans les étapes précoces du cycle viral (Violot S, Hong SS et al. 2003). Par contre, Rakotobe et ses collègues suggèrent que EED exercerait une activité antivirale durant la phase tardive du cycle de réplication, au moment du rassemblement de la nouvelle particule virale, et que son interaction avec Nef semble inverser cet effet négatif (Rakotobe D, Tardy JC et al. 2007).

*c) RAD18*

Rad18 appartient à la machinerie de réparation de l'ADN (Broomfield S, Hryciw T et al. 2001; Inagaki A, Sleddens-Linkels E et al. 2011). Une interaction entre Rad18 et l'IN de VIH-1 a été montrée par des études de co-immunoprécipitation et de co-localisation (Mulder LC, Chakrabarti LA et al. 2002). L'interaction entre Rad18 et l'IN pourrait permettre de stabiliser l'IN (Mulder LC, Chakrabarti LA et al. 2002).

Par ailleurs, Rad18 pourrait inhiber l'infection virale d'une façon non-spécifique en interagissant avec l'ADN viral (Lloyd AG, Tateishi S et al. 2006).

*d) RAD51*

Une interaction physique entre l'IN de VIH-1 et RAD51 a été démontrée au départ par ELISA. *In vitro*, l'addition de RAD51 affecte négativement la réaction d'intégration concertée catalysée par l'IN, suggérant un rôle de répression de l'intégration par RAD51 (Desfarges S, San Filippo J et al. 2006). RAD51 est supposée agir soit par un effet direct sur l'IN, soit par compétition avec l'IN pour son substrat.

*e) Daxx*

Daxx (Death associated protein) a été identifiée comme partenaire de l'IN d'ASV par double hybride puis par co-immunoprécipitation (Greger JG, Katz RA et al. 2005), et récemment retrouvée comme partenaire de l'IN de VIH-1 (Huang L, Xu GL et al. 2008). Daxx n'est pas requise pour l'intégration mais serait impliquée dans la régulation négative de l'expression des gènes viraux via le recrutement des HDACs.

*f) HSP60*

Les protéines de choc thermique sont connues pour leur inductibilité par la chaleur et sont généralement responsables de la protection, le maintien et la régulation des fonctions des protéines auxquelles elles sont associées. L'interaction entre HSP60 (Heat Shock Protein60) et l'IN de VIH-1 a été démontrée par chromatographie d'affinité (Parissi V, Calmels C et al. 2001). Les mêmes auteurs ont montré que HSP60 semble stimuler les réactions de clivage des extrémités 3' et transfert de brins réalisées par l'IN de VIH-1 *in vitro* (Parissi V, Calmels C et al. 2001). HSP60 pourrait donc agir comme un chaperon afin de protéger l'IN contre une

inactivation par la chaleur, ou elle pourrait agir sur la conformation de l'IN pour induire ou stabiliser une forme active de la protéine.

### **III- Les partenaires cellulaires de l'IN de MLV**

Des travaux menés par Studamire et ses collègues ont mis en évidence de nombreux partenaires de l'IN de MLV (Studamire B and Goff SP 2008). Certains sont des facteurs de réparation de l'ADN, d'autres sont des facteurs de liaison à l'ARN, ou à la chromatine. Parmi ces derniers, le facteur de transcription Brd2 fera l'objet de notre étude.

#### *A. Les facteurs de réparation de l'ADN*

La première protéine de réparation de l'ADN isolée est la Ku70/XRCC6, une sous-unité de 70 KDa de l'hétérodimère Ku (Ku70/Ku80 thyroid autoantigen). L'hétérodimère Ku est impliqué dans la voie NHEJ de réparation de l'ADN (Hartley KO, Gell D et al. 1995), et participe avec TRF2 (Telomere repeat factor 2) pour inhiber la recombinaison homologue des télomères entre deux chromatides sœurs (Celli GB, Denchi EL et al. 2006).

Fen-1 (Flap endonuclease-1), ou Rad27 ou RTH1 (RAD two homolog-1) est endo/exonucléase qui agit en maintenant la stabilité du génome, dans voie NHEJ, et dans résolution des fragments Okazaki durant la synthèse d'ADN (Liu Y, Kao HI et al. 2004). Notons que le domaine C-terminal de Fen-1 interagit avec p300, qui joue un rôle dans l'intégration rétrovirale, et qui acétyle Fen-1 (Hasan S, Stucki M et al. 2001).

#### *B. Les facteurs de liaison à l'ARN*

Parmi les protéines de liaison à l'ARN, qui ont été isolées par double hybride comme partenaires de l'IN de MLV, on retrouve SMN (Survival motor neuron), U2AF<sup>26</sup> (U2snRNP associated factor), U5 snRNP, et SF3a (Splicing factor 3b).

snRNPs (Spliceosomal small ribonucleoproteins) constituent les composants principaux de la machinerie d'épissage de l'ARN, et chaque snRNP est constitué d'un ou de deux snRNAs (small nuclear RNAs) liés à un ensemble de protéines, dites les protéines Sm (Battle DJ, Kasim M et al. 2006). L'assemblage de ces protéines aux snRNAs se fait grâce au complexe SMN. Notons que le complexe SMN interagit directement avec les protéines Gemin, dont Gemin2 un partenaire de l'IN de VIH-1. Le complexe SMN interagit également avec snRNPs U1, U2 et U5, des composants du spliceosome.

L'identification des facteurs SMN, U2AF<sup>26</sup>, U5snRNP et SF3a par double hybride comme partenaires de l'IN de MLV, permet de suggérer qu'il existe une interaction entre l'ARN viral et le spliceosome.

### *C. Les facteurs de liaison à la chromatine*

Une catégorie de protéines isolées par double hybride qui est particulièrement intéressante concerne les facteurs de liaison à la chromatine et les régulateurs de la transcription. Une première protéine isolée est Enx-1/Ezh2 (Enhancer of Zeste homolog 1), un membre du complexe PRC2 (Polycomb repressive complex 2) dont EED fait partie (partenaire de l'IN du VIH-1). En effet, Enx-1 et EED interagissent *in vitro* et *in vivo* (van Lohuizen M, Tijms M et al. 1998). Bien qu'il n'existe aucune donnée qui prouve que l'IN de MLV interagit avec EED, l'isolation de Enx-1 comme partenaire de l'IN de MLV, et la validation de son interaction avec l'IN de VIH-1, permettent de suggérer que cette protéine pourrait jouer un rôle dans l'intégration rétrovirale.

AF9 (Acute lymphocytic leukemia gene 1 fused from chromosome 9), également connue sous le nom Mllt3 (mixed lineage leukemia translocated to chromosome 3 homolog), a été isolée comme partenaire de l'IN de MLV et validée comme partenaire de l'IN de VIH-1. La fonction de AF9 est inconnue, mais elle est supposée être un régulateur de la transcription dû au fait qu'elle contient un domaine riche en sérine et proline, ainsi qu'un signal de localisation nucléaire. Notons que le domaine C-terminal de AF9 interagit avec Pc3 (Drosophila Polycomb group protein) murine et humaine, mais également avec BcoR (BCL6 corepressor), sachant que Pc3 et BcoR sont impliqués dans la répression de la transcription (Hemenway CS, de Erkenez AC et al. 2001; Srinivasan RS, de Erkenez AC et al. 2003).

B-ATF est un membre de la superfamille de facteurs de transcriptions AP-1/ATF (Dorsey MJ, Tae HJ et al. 1995). Ce facteur est exprimé dans les cellules et tissus hématopoïétiques (Hemenway CS, de Erkenez AC et al. 2001). B-ATF ne dimérise pas avec Fos, mais forme des hétérodimères avec la famille de protéines Jun (c-Jun, JunD et JunB) pour se fixer aux sites d'ADN consensus d'AP-1, et agit comme un régulateur négatif de la transcription d'AP-1. Le clone de B-ATF isolé par double hybride n'interagit pas avec l'IN de VIH-1.

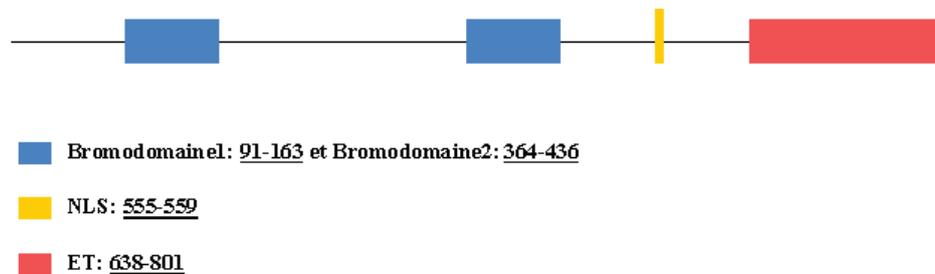
#### *D. Brd2, partenaire potentiel de l'IN de MLV et de VIH-1*

Brd2/Ring3 est un régulateur de transcription appartenant à la famille de protéine BET (Bromodomains and Extra Terminal domain). Cette famille de régulateurs sont impliqués dans des processus cellulaires larges comprenant la réplication, la transcription, l'inactivation de gènes et le remodelage de la chromatine (Wu SY and Chiang CM 2007; LeRoy G, Rickards B et al. 2008; Denis GV, Nikolajczyk BS et al. 2010; Mantovani F, Drost J et al. 2010; Weidner-Glunde M and Ottinger M 2010 ). Les membres de cette famille sont caractérisés par la présence d'un ou de deux bromodomaines en tandem (BD1 et BD2) en N-terminal, impliqués dans la liaison aux histones acétylés, et d'un domaine extraterminal (ET) impliqué dans l'interaction protéine-protéine (figure 18). Le bromodomaine est un motif de 110 acides aminés qui se lie spécifiquement à la lysine 12 acétylée de l'histone H4 (Kanno T, Kanno Y et al. 2004). Des études structurales ont permis de caractériser cette interaction pour la protéine Brd2 (Nakamura Y, Umehara T et al. 2007; Umehara T, Nakamura Y et al. 2010). En effet, la structure cristallographique du bromodomaine N-terminal (BD1, résidus 74-194) de la protéine Brd2 humaine montre que BD1 forme un dimère, et que le site de liaison à l'acétyllysine formé à l'interface de ce dimère permet l'interaction avec l'histone H4 acétylé. Le bromodomaine interagit également avec la lysine 382 acétylée du p53, une protéine impliquée dans la suppression des tumeurs (Eichenbaum KD, Rodríguez Y et al. 2010).

*Brd2/ RING3/ fsrg1* a été initialement identifié dans la région MHC (Major Histocompatibility Complex) classe II, sur le chromosome humain 6p21.3, relié au gène *fsrg* de la drosophile (*Drosophila female sterile gene*). Ce gène a été initialement nommé RING3 (really intresting new gene-3), puis nommé plus tard BRD2. Depuis, 13 gènes homologues ont été identifiés, chez la souris, le rat, le singe, le chimpanzé, la grenouille, le poisson zèbre, le saumon, le chien, les bovins, et les aviaires.

La protéine Brd2 (RING3) est une sérine/thréonine kinase localisée dans le noyau. En effet, Brd2 contient un domaine « kinase-like » entre ces deux bromodomaines, responsable de son activité kinase en réponse à un signal mitogénique. Elle possède également une fonction de régulation de la transcription, et peut aussi s'associer à la chromatine acétylée durant la mitose, alors que les autres membres de la famille BET se dissocient du chromosome durant la mitose. Ainsi Brd2 peut activer la transcription. Elle s'associe aux complexes contenant E2F (E2 promoter-binding factor) et recrute des HAT et des facteurs de remodelage de la chromatine (Denis GV, McComb ME et al. 2006) au promoteur de la

cycline A (Sinha A, Faller DV et al. 2005), activant ainsi la transcription (LeRoy G, Rickards B et al. 2008). Une étude récente a montré que Brd2 peut également recruter des co-répresseurs de la transcription ; ainsi, Brd2 co-réprime le gène PPAR $\gamma$  (peroxisome-proliferator activated receptor gamma) dans les adipocytes et le gène insuline dans les cellules pancréatiques  $\beta$  (Wang F 2009).



**Figure 18. Organisation en domaines de Brd2.** La protéine Brd2 est constituée de 2 bromodomains et d'un domaine extra-terminal. Elle contient également une séquence de localisation nucléaire représentée sur ce schéma.

Plusieurs rôles ont été décrits pour la protéine Brd2 dans la cellule, on cite :

- Rôle dans l'obésité et le diabète

Des expériences de délétion du gène *Brd2* ont été, au début, difficiles vu que la protéine est essentielle pour la croissance cellulaire, et sa délétion cause une létalité (Gyuris A, Donovan DJ et al. 2009; Shang E, Wang X et al. 2009). Or, il y a eu récemment une découverte accidentelle de souris résistantes qui possèdent un gène *lacZ* inséré en 5' des régions contrôlant le gène *brd2*. Cette mutation réduit mais n'élimine pas l'expression de Brd2 et préserve donc la survie de la souris. Par contre, les souris deviennent extrêmement obèses et présentent une tolérance au glucose. Ces observations ont permis d'identifier un rôle essentiel de Brd2 dans la fonction des cellules pancréatiques  $\beta$ , et dans l'adipogenèse (Wang F 2009).

- Rôle dans le cycle cellulaire et dans le cancer

Brd2 est impliquée dans le cancer humain. Une surexpression de cette protéine co-active le promoteur de la cycline A, et provoque une entrée précoce dans la phase S durant la progression du cycle cellulaire (Sinha A, Faller DV et al. 2005). Une déstabilisation du cycle

cellulaire entraîne un développement de lymphomes aux cellules B et de leucémies dans les lymphocytes de souris (Greenwald RJ, Tumang JR et al. 2004).

#### - Rôle dans l'inflammation

Chez l'homme, le gène *Brd2* se trouve sur le chromosome 6 (bande 6p21.3) du complexe majeur d'histocompatibilité. Cette région est associée aux gènes du système immunogénétique HLA (« human leucocyte antigen »), impliqués dans diverses conditions inflammatoires (de Bakker PI, McVean G et al. 2006).

Chez la souris, *Brd2* est considérée comme un régulateur critique de l'inflammation des adipocytes (Wang F 2009) et des macrophages infiltrant le tissu adipeux (Belkina AC and Denis GV 2010). D'autre part, un faible taux de *Brd2* est impliqué dans le mécanisme de protection des individus MHO (« metabolically healthy despite obesity ») des complications inflammatoires liées à leurs obésités (Belkina AC and Denis GV 2010).

#### - Rôle dans la réplication virale

A part son rôle potentiel dans la réplication des rétrovirus (MLV et VIH-1), *Brd2* est également impliquée dans la réplication du Kaposi's sarcoma-associated herpesvirus (KSHV). Durant une infection latente, le génome d'un herpesvirus est maintenu sous forme d'épisomes circulaires stables dans le noyau des cellules infectées. La transmission du génome viral durant la mitose est assurée par une interaction du génome des épisomes avec la protéine LANA-1 (Latency-associated nuclear antigen 1), un des produits des gènes de latence du virus. LANA-1 semble également interagir avec *Brd2* pour attacher le génome de KSHV aux chromosomes mitotiques, démontrant un rôle de *Brd2* dans la réplication de ce virus (Viejo-Borbolla A, Ottinger M et al. 2005).

Le fait que *Brd2* soit capable de se fixer à la chromatine et étant un partenaire cellulaire de l'intégrase, permettent de suggérer que cette protéine pourrait jouer un rôle dans l'intégration des rétrovirus, notamment dans la sélection des sites d'intégration. Notons que des données précédentes ont montré que des protéines cellulaires peuvent être impliquées dans le ciblage des sites d'intégration. C'est le cas de LEDGF/p75 dans l'intégration du VIH-1. C'est également le cas des rétrotransposons Ty1, Ty3 et Ty5. En effet, le ciblage des sites d'initiation de la transcription du gène tRNA par Ty3 implique une interaction avec TFIIB et

TFIIIC (Chalker DL and Sandmeyer SB 1992). Un autre exemple est l'intégration de Ty5 dans les régions d'hétérochromatine qui se perd suite à une mutation du facteur Sir4p indiquant que le ciblage des sites d'intégration peut être contrôlé par des régulateurs de la transcription (Sandmeyer S 2003).

## **OBJECTIFS**

Les informations acquises sur l'IN des rétrovirus durant ces dernières années ont permis d'avancer significativement la compréhension de son rôle dans le cycle viral. D'importantes questions ont été soulevées concernant l'organisation et la fonction des différents domaines de l'IN, ses activités catalytiques, son état oligomérique, ses partenaires cellulaires et viraux, etc. La résolution de la structure des domaines de l'IN et récemment de la protéine entière a fourni également des modèles qui vont faciliter la conception rationnelle d'inhibiteurs de cette enzyme. Cependant, il reste des questions largement inexplorées sur l'identité et l'importance des partenaires cellulaires de l'IN.

Le présent travail a donc pour objectif (i) d'optimiser les protocoles de production de protéines intégrases pures et actives, (ii) d'identifier des partenaires cellulaires de l'IN. Deux modèles d'enzyme rétrovirales ont été utilisés dans cette étude :

- L'IN du virus RAV-1, un rétrovirus exogène aviaire du genre des alpharétrovirus appartenant au sous-groupe A de la famille des ASLV. Comme nous avons eu l'occasion de le voir, l'IN d'ASLV est largement étudiée soit au niveau structural ou fonctionnel. Dans notre équipe à Lyon, les activités catalytiques de l'IN RAV-1 *in vitro* ont été explorées afin de mieux comprendre le fonctionnement de l'enzyme *in vivo* (Moreau K, Faure C et al. 2003; Moreau K, Faure C et al. 2004; Moreau K 2009; Ballandras A, Moreau K et al. 2011). Or, à ce jour, il n'existe toujours pas de données suffisantes sur les partenaires cellulaires de cette enzyme.

- L'IN de PERV A/C, un rétrovirus endogène porcin du genre gammarétrovirus. Concernant la composition de ce virus, les données en rapport avec les différentes enzymes virales sont limitées. Une étude récente, menée par notre équipe à Ploufragan, a montré que les virus PERV et MLV possèdent des similarités dans le mécanisme de sélection des sites cellulaires d'intégration de l'ADN viral (Moalic Y, Blanchard Y et al. 2006; Moalic Y, Félix H et al. 2009). Ce résultat suggère que l'IN PERV devrait avoir des caractéristiques communes avec celle de MLV, en particulier au niveau des cofacteurs cellulaires impliqués dans le choix de ces sites. Cependant, aucune information sur sa structure, ses activités, ou ses cofacteurs n'a été obtenue jusqu'à présent.

Afin de produire des intégrases pures et actives, nous avons choisi de comparer les tests d'activités catalytiques *in vitro* des intégrases produites à partir du système bactérie ou du système baculovirus / cellule SF9. Ce dernier devrait conférer aux protéines des modifications

post-traductionnelles durant la production, nécessaires pour certaines de leurs activités, et dont le système bactérien n'est pas capable d'accomplir.

D'autre part, pour identifier des co-facteurs cellulaires de ces deux enzymes, nous nous sommes basés sur les résultats obtenus par un double hybride réalisé pour l'IN de JSRV par l'équipe de Vincent Lotto (U851- SFR BioSciences Gerland - Lyon Sud (UMS3444/US8)). Notons que le JSRV est un bêtarétrovirus, responsable d'un adénocarcinome pulmonaire chez le mouton. Cette étude a permis de révéler 2 partenaires cellulaires de l'IN de JSRV, identifiés comme Brd2 et TAB182 (182 kDa Tankyrase-1-Binding protein). Les résultats obtenus par double hybride, mené par Studamire et ses collègues, ont montré que Brd2 est également un partenaire de l'IN de MLV (Studamire B and Goff SP 2008). Or, le fait que Brd2 soit un régulateur de la transcription capable de se fixer aux histones acétylés nous a laissé supposer qu'il pourrait avoir un rôle dans le ciblage de sites spécifiques dans l'ADN cellulaire durant le processus d'intégration de l'ADN viral. Ceci nous a encouragés à l'étudier plus précisément et à vérifier son interaction avec nos deux modèles d'INs.

Ainsi, ce travail de thèse a visé à valider et caractériser les interactions potentielles entre la protéine cellulaire (Brd2) et les INs de PERV et d'ASLV. La comparaison entre ces deux modèles devrait permettre de déterminer si ce co-facteur cellulaire est un partenaire commun à toutes les INs, ou propre à certains genres viraux, leur procurant des caractéristiques spécifiques à leur mécanisme d'intégration.

## **MATERIELS ET METHODES**

## I- Construction des vecteurs d'expression de l'IN PERV et ses différents domaines

Nom du plasmide	Construction	Amorces
pGVB3040	N-ter domain of PERV IN dans pVL1392	OGVB3050-F (GCGGCCGCACATGGCACATCACCATCACCATCACCATAGAAATGCCAAAG) OGVB3071-R (GAATTCTAAGCATTAAACCAGCTGGCAGGGC)
pGVB3041	Core domain of PERV IN dans pVL1392	OGVB3070-F(GCGGCCGCCATGGCACATCACCATCACCATCACCATCACAATCCTTCCAGAATGCCTCC) OGVB3069-R (GAATTCTATACAGAAGCAATTTCTACCAATGG)
pGVB3038	C-ter domain of PERV IN dans pVL1392	OGVB3068-F (GCGGCCGCCATGGCACATCACCATCACCATCACCATCAGAAATGCTTCTGTACATAGTGC) OGVB3049-R (GAATTCTTAACCTTGAGGAGTTATTGACAGAG)
pGVB3039	N-ter/core domains of PERV IN dans pVL1392	OGVB3050-F (GCGGCCGCACATGGCACATCACCATCACCATCACCATAGAAATGCCAAAG) OGVB3069-R (GAATTCTATACAGAAGCAATTTCTACCAATGG)
pGVB3037	Core/C-ter domains of PERV IN dans pVL1392	OGVB3070-F(GCGGCCGCCATGGCACATCACCATCACCATCACCATCACAATCCTTCCAGAATGCCTCC) OGVB3049-R (GAATTCTTAACCTTGAGGAGTTATTGACAGAG)
pGVB3030	WT PERV IN dans pVL1392	OGVB3050-F (GCGGCCGCACATGGCACATCACCATCACCATCACCATAGAAATGCCAAAG) OGVB3049-R (GAATTCTTAACCTTGAGGAGTTATTGACAGAG)
pGVB3045	WT PERV IN dans pET14b	OGVB3074-F (CATATGATAGAAATGCCAAAGCCCCAGAAC) OGVB3074-R (CATATGTTAACCTTGAGGAGTTATTGACAGAG)

**Tableau 2. Amorces utilisées pour la construction des différents plasmides.** Séquences nucléotidiques des différentes amorces utilisées.

### *A. PCR*

La séquence de l'intégrase de PERV a été amplifiée en PCR en utilisant le plasmide A14/220 comme matrice. Pour ce fait, on mélange 10 ng/μl de l'ADN matrice avec 1μl de chacune des amorces à 10μM nécessaires pour ces constructions (voir tableau 2), ainsi qu'1μl de dNTP à 10mM, 5 μl de tampon Thermopol Reaction buffer 10X, 1 unité de Vent DNA Polymerase (Biolabs), dans un volume final de 50 μl d'eau. La PCR est réalisée sur un thermocycler (Eppendorf) selon les conditions suivantes : 95 °C pendant 5 minutes, puis 30 cycles à 95 °C pour 30 secondes, 55 °C pour 60 secondes, et 72 °C pour 1.5 minutes, et enfin 72 °C pendant 7 minutes.

### *B. Digestion et ligation*

Pour la construction des plasmides pGVB3030, pGVB3037 pGVB3038, pGVB3039, pGVB3040, pGVB3041, les produits de PCR sont digérés pendant 1 heure à 37°C par les enzymes de restriction NotI et EcoRI en quantités recommandées par le fournisseur (Fermentas), puis migrés à 100volts sur gel d'agarose 1% (1g d'agarose dans 100 ml tampon Tris/Borate/EDTA). Le fragment correspondant à la taille attendue est extrait du gel par « MinElute Qiaquick Gel Extraction Kit » (Qiagen), et élué 2 fois dans 10μl de tampon

EB. Les produits de digestion (6µl) sont incubés durant la nuit à 16°C avec 2µl du plasmide PVL1392 (BD baculogold), en présence de 100 unités de l'enzyme T4 DNA ligase (Biolabs).

Pour la construction du plasmide pGVB3045, les produits de PCR sont digérés par les enzymes de restriction NdeI et HindIII (Fermentas) puis insérés dans le plasmide pET14b (Invitrogen) qui contient une étiquette 6x histidine en N-terminal.

### *C. Transformation de bactéries*

Des bactéries TOP10 sont ensuite transformées durant 15 minutes dans la glace avec la totalité du produit de ligation. La transformation est réalisée par un choc thermique à 42°C pendant 30 secondes. 250µl de SOC sont ensuite ajoutés et la réaction est incubée pendant 1 heure à 37°C sous agitation (200rpm). 100µl sont ensuite étalés sur boîte de gélose en présence de l'ampicilline (50 µg/ml) et incubés durant la nuit à 37°C. Des colonies sont sélectionnées et les plasmides sont préparés en utilisant « QIAprep Spin Miniprep Kit » (Qiagen).

### *D. Séquençage*

#### *a) Principe du séquençage*

Le séquençage de l'ADN a été effectué au laboratoire selon la méthode "dideoxy chain terminator" mise au point par Sanger. Elle nécessite une PCR en présence de didéoxynucléotides triphosphates (ddNTPs) dépourvus de groupements hydroxyles en 3', et qui stoppent l'élongation de la molécule d'ADN par les ADN polymérases. Les ddNTPs utilisés sont marqués chacun avec un fluorophore spécifique. Ainsi, après dénaturation, lorsque les fragments sont séparés par électrophorèse capillaire, une analyse automatique par un système laser permet d'identifier le fluorophore terminateur des fragments dont la taille varie en fonction de la position d'intégration du ddNTP. Ceci permet d'obtenir la séquence de la portion d'ADN amplifiée.

#### *b) Protocole*

Les plasmides d'une colonie positive pour l'insert sont amplifiés et purifiés en suivant les instructions du fournisseur du kit " MinElute Qiaquick Gel Extraction Kit " commercialisé par Qiagen. Une mesure photométrique à 260 nm a permis de quantifier l'ADN plasmidique

purifié. Pour 20  $\mu\text{L}$  final, on a utilisé 4  $\mu\text{L}$  de mélange réactionnel contenant une ADN polymérase et un mélange de dNTP et de ddNTP, ces derniers étant couplés chacun à un chromophore différent. Une quantité de 300 ng de plasmide a été utilisée comme matrice et 10 pmol des amorces Polyhedrin fwd et rev. Le programme PCR a consisté en 25 cycles à 96°C pendant 10 secondes, 50°C pendant 5 secondes et 60°C pendant 4 minutes. Les produits d'extension ont été purifiés par précipitation à l'éthanol. Les séquences ont été réalisées à l'aide d'un appareil de modèle ABI Prism 310.

## **II- Production des intégrases d'ASLV et de PERV**

### *A. Production en bactérie*

Les bactéries BL21 (DE3) pLysS *E.Coli* (Invitrogen) sont transformées par les plasmides (pGVB3045 et pET14b-INASLV) contenant les gènes d'intérêt. Ces bactéries possèdent le lysogène  $\lambda$  DE3 qui contient le gène *LacI*, le gène de la T7 RNA polymérase sous le contrôle du promoteur lac UV5. Ce promoteur est inductible par l'IPTG. Ces bactéries possèdent également le plasmide pLysS qui produit le lysozyme T7 permettant la réduction du niveau basal d'expression du gène d'intérêt en absence d'IPTG.

Protocole :

Les bactéries sont incubées avec 1ng de plasmide pendant 30 minutes dans la glace. La transformation se fait par choc thermique, réalisé à 42°C pendant 30 secondes, suivi de 2 minutes d'incubation dans la glace. On rajoute 900  $\mu$ l de milieu SOC (Invitrogen) et on incube 1 heure à 37°C sous agitation. Les clones bactériens sont sélectionnés sur boîte gélosée en présence d'ampicilline à 50  $\mu$ g/ml. Les clones sélectionnés sont mis en pré-culture dans 3 ml de milieu LB contenant de l'ampicilline à 50  $\mu$ g/ml, pendant une nuit, à 37°C.

Le lendemain, la pré-culture est utilisée pour ensemercer 1 L de milieu LB. Lorsque la DO (densité optique) à 600 nm atteint 0.6, l'expression des protéines est alors induite par l'addition de 1 mM final d'IPTG. Les cultures bactériennes sont incubées pendant 3 heures à 37°C sous agitation. Les bactéries sont enfin centrifugées à 6000 g, pendant 15 minutes à 4°C. Le culot bactérien est conservé à -80°C jusqu'à son utilisation.

### *B. Production en baculovirus/cellule d'insecte*

Le principe de production en système baculovirus / cellules SF9 repose sur la co-transfection des cellules d'insecte par un plasmide de transfert recombinant (pGVB 3030 et pGVB3033) et l'ADN du baculovirus sauvage linéarisé. Le gène d'intérêt et le gène *lacZ* présents sur le plasmide de transfert sont encadrés de séquences de recombinaison identiques à celles situées sur l'ADN du baculovirus sauvage. Il se produit alors une recombinaison homologue entre le plasmide recombinant et le génome du baculovirus ce qui permet la production de baculovirus recombinants. De plus, le gène *lacZ* étant transféré en même temps que le gène d'intérêt, il est possible de détecter les cellules infectées par le baculovirus recombinant : elles se colorent en bleu en présence de X-Gal. Les virus recombinants sont

sélectionnés et utilisés pour infecter des cellules afin de produire en grande quantité la protéine d'intérêt. Les cellules utilisées sont les SF9, qui proviennent du tissu ovarien du lepidoptère *Spodoptera frugiperda*.

Protocole :

Les cellules SF9 ( $2 \times 10^6$  cellules), sont transfectées avec 0.5  $\mu\text{g}$  plasmide BD BaculoGold DNA linéarisé (BD Biosciences) et 4.5 $\mu\text{g}$  du plasmide recombinant, en présence de 15  $\mu\text{l}$  Cellfectine (Invitrogen), dans un volume de 2 ml de milieu TG1 sans antibiotiques. Les cellules sont incubées pendant 5 heures à 28°C, puis le milieu est remplacé avec 4 ml de TG3 additionné de 5% SVF. Le surnageant est récupéré 5 jours après la transfection, puis centrifugé pendant 10 minutes à 2200 rpm. Le surnageant est alors utilisé pour l'infection de  $4 \times 10^6$  cellules en culture dans un spinner de 250 ml.

Le culot d'une culture de  $1 \times 10^6$  cellules est resuspendu dans 500  $\mu\text{l}$  de tampon "insect cell lysis buffer" (BD BaculoGold), en présence de 1% (v/v) de détergent Nonidet P-40 (NP40). Les échantillons sont ensuite soniqués 3 fois 20 secondes et centrifugés à 13000 rpm pendant 20 minutes à 4°C. L'expression des protéines est contrôlée par western blot en utilisant un anticorps anti-histidine (Pierce). Le reste du culot est conservé à -80°C.

### **III- Purification des intégrases**

#### *A. Les solutions*

- Tampon A (de lyse) : 20 mM Tris HCl pH8, 500 mM NaCl, 10 mM Imidazole, 5 mM  $\beta$ -mercaptoéthanol

- Tampon B1 (d'équilibre) : 20 mM Tris HCl pH8, 500 mM NaCl, 2.5 mM Imidazole, 5 mM  $\beta$ -mercaptoéthanol

- Tampon B2 (de lavage) : 20 mM Tris HCl pH8, 1 M NaCl, 25 mM Imidazole, 5 mM  $\beta$ -mercaptoéthanol

- Tampon C (d'élution) : 20 mM Tris HCl pH8, 500 mM NaCl, 1 M Imidazole, 5 mM  $\beta$ -mercaptoéthanol

- Tampon de dialyse : 20 mM Tris HCl pH8, 500 mM NaCl, 5 mM  $\beta$ -mercaptoéthanol, 10  $\mu\text{M}$   $\text{ZnCl}_2$ .

## *B. Le protocole*

Les protéines intégrases de PERV et d'ASLV, produites en bactéries et SF9, sont purifiées par chromatographie d'affinité sur colonne de nickel grâce à l'affinité de l'étiquette histidine présente à l'extrémité N-terminale de la protéine pour les ions nickel. Ces purifications sont réalisées sur un appareil AKTA<sub>FPLC</sub> (Amersham). Pour cela, les culots bactériens ou cellulaires sont repris dans 10 ml de tampon de lyse A auquel sont ajoutés des inhibiteurs de protéase (complete Mini EDTA free (Roche)). Les culots sont incubés 1 heure dans la glace. Au bout de 30 minutes, 10µg/ml de DNase et de RNase sont ajoutés ainsi qu'1.5 mg/ml de lysosyme. Ensuite, une sonication de 5 fois 30 secondes est réalisée pour les protéines produites an bactérie, et 3 fois 20 secondes pour les protéines produites en cellules SF9. Les lysats bactériens sont ensuite centrifugés pendant 30 minutes, à 10 000 rpm, 4°C, alors que les lysats des cellules SF9 sont centrifugés pendant 20 minutes à 13 000 rpm, 4°C.

Le surnageant est alors injecté à 0.5ml/min sur une colonne de chromatographie d'affinité « Hitrap IMAC HP columns » (GE Healthcare). La colonne est préalablement lavée avec 3 ml d'eau distillée, puis 3 ml de tampon B1, et chargée avec une solution de NiSO<sub>4</sub> à 0.1M. La colonne est ensuite lavée avec 10 ml de tampon de lavage B2, puis par 5 ml de tampon B1, pour éliminer les protéines non fixées. L'élution se fait grâce à un gradient d'imidazole de 5 mM à 1 M, entraînant le décrochage de la protéine. Les protéines sont collectées par fraction de 500 µl et sont analysées sur gel SDS PAGE 10% suivi d'une coloration au bleu de coomassie pendant 20min afin de visualiser les protéines. Les fractions pures sont conservées et dialysées pendant une nuit à 4°C contre un litre de tampon de dialyse. Les protéines sont enfin conservées à -80°C jusqu'à leur utilisation. Les intégrases d'ASLV sont dialysées en absence de ZnCl<sub>2</sub> et sont indiquées par ASLV IN<sub>bact</sub> ou ASLV IN<sub>SF9</sub>. Pour les intégrases de PERV, une partie est dialysée en présence de 10µM ZnCl<sub>2</sub> et sont indiquées par PERV IN<sub>bact-zc</sub> ou PERV IN<sub>SF9</sub> et une partie de protéines sont dialysées en absence de ZnCl<sub>2</sub> et sont indiquées par PERV IN<sub>bact</sub>.

## **IV- Analyse et dosage des fractions d'IN purifiées sur gel SDS-PAGE**

La pureté des protéines est contrôlée sur gel SDS PAGE 12%. Le dosage des protéines pures est également réalisé sur gel SDS PAGE 12%, par comparaison des échantillons d'intégrases pures à une gamme de protéines BSA pures de concentrations variant entre 100 et 800 ng/µl.

1- Les fractions pures sont diluées dans le tampon de dépôt (100 mM Tris-HCl pH8, 10 % SDS, 5 % glycérol, 300µg/ml bleu de bromophénol, 100 mM DTT), puis dénaturées 5 minutes à 95°C.

2- Les protéines sont déposées sur un gel SDS-PAGE 12%. Pour cela, on prépare un petit gel de résolution de 10ml (3.3 ml d'acryl-bisacrylamide 30%, 2.5 ml de Tris- HCl à 1.5 M, pH 8.8, 0.1 ml de SDS 10%, 0.1 ml d'APS 10% et 0.004 ml de TEMED), puis un gel de concentration de 4ml (0.67 ml d'acryl-bisacrylamide 30%, 0.5 ml Tris-HCl à 0.5 M, pH6.8, 0.04 ml de SDS 10%, 0.04 ml d'APS 10 %, 0.004 ml TEMED). On coule les gels et on ajoute le peigne.

3- La migration est effectuée dans le tampon de migration (25 mM Tris, 192 mM glycine, 0.1% SDS), à 200 volts.

4- Le gel est ensuite coloré pendant 60 minutes dans une solution PageBlue™ Protein Staining Solution (Fermentas, Ref : R0571), puis décoloré dans plusieurs bains de 5 minutes dans de l'eau distillée, afin de visualiser les protéines.

#### **V- Analyse des formes oligomériques de l'intégrase par fixation chimique**

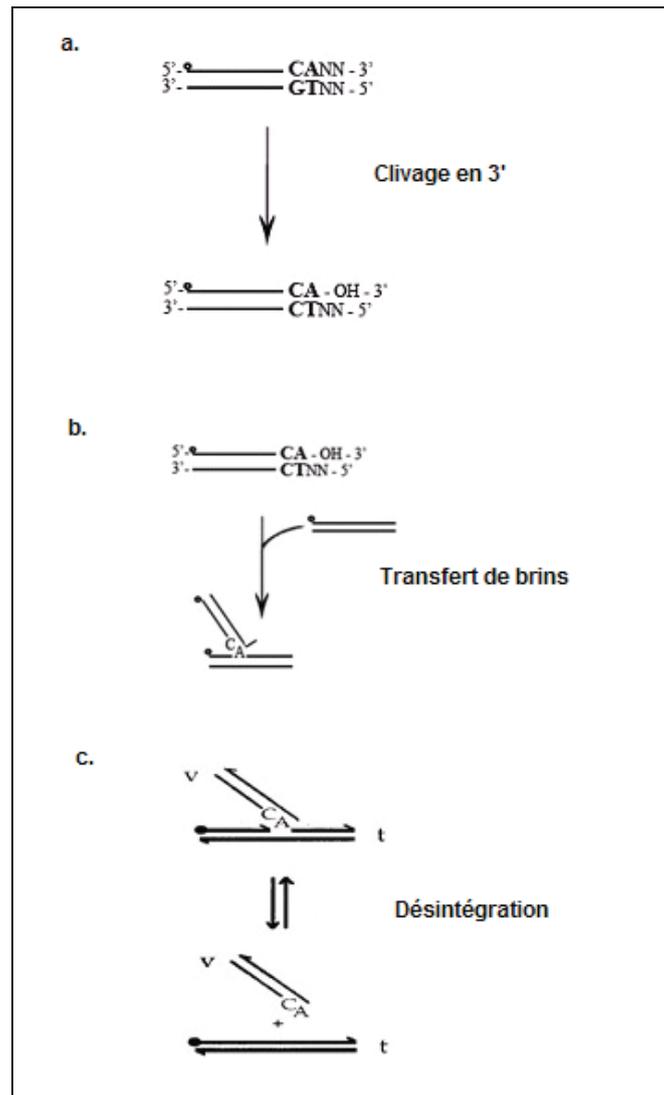
Afin de déterminer sous quelle(s) forme(s) multimérique(s) se trouve l'intégrase en solution, différentes formes sont alors fixées par un agent chimique, le BS<sup>3</sup> [Bis (sulfosuccinimidyl) suberate]. Pour cela, les protéines IN de PERV, et IN d'ASLV (2µg) sont mélangées avec 2 mM de BS<sup>3</sup> et incubées pendant 30 minutes à température ambiante dans un tampon contenant 100 mM d'Hepes à pH8 et 600 mM de NaCl. Les réactions sont arrêtées par addition de la solution d'arrêt (250 mM tris-Hcl pH7.5, 30 mM glycine) puis incubées 15 minutes à température ambiante.

Les différentes formes oligomériques sont séparées sur gel d'acrylamide 8% (préparé comme décrit ci-dessus). Le gel est ensuite transféré pendant 1 heure à 100 volts sur une membrane nitrocellulose. La membrane est colorée au rougeponceau pendant quelques minutes puis rincée à l'eau déminéralisée. Elle est ensuite bloquée pendant 1 heure dans une solution de TBS-T (Tris-buffered saline, 0.05% Tween 20) contenant 25mg/ml BSA (Bovine Serum Albumin), lavée 4 fois durant 5 minutes dans du TBS-T, puis incubée dans 10 ml de TBST contenant 1/5000 d'anticorps anti-histidine-HRP (Thermo Scientific). Les différentes formes oligomériques sont enfin détectées en utilisant le kit ECL (Thermo scientific) pour la révélation.

## VI- Tests des activités catalytiques de l'intégrase

### A. Principe

Les activités de clivage, transfert de brins et désintégration de l'intégrase sont testées en présence d'oligonucléotides mimant la séquence virale U5. Ces oligonucléotides sont marqués par un phosphate radioactif au niveau de leur extrémité 5'. Les produits issus de ces réactions sont observés sur gel dénaturant (Figure 19).



**Figure 19. Principe des activités catalytiques *in vitro* de l'intégrase.** Pour la réaction de clivage en 3' (a), les oligonucléotides mimant la séquence U5 virale non clivée sont mis en présence de l'intégrase. Pour la réaction de transfert de brins (b), l'intégrase est mise en présence des oligonucléotides déjà pré-clivés présentant donc une extrémité CA-OH en 3'. Pour la réaction de désintégration (c), l'intégrase est mise en contact avec les produits de la réaction de transfert de brins.

## *B. Protocole*

### *a) Marquage radioactifs des oligonucléotides*

Afin de marquer les oligonucléotides, 10 pmol de PERV-U5A (5'-CTTTTTGCTGGTCTTACATT -3'), PERV-U5AC (5'-CTTTTTGCTGGTCTTACA -3'), PERV-D (5'-TGTAAGATCTTACAGGCCCTTGGGCCGGCGCTTGCGCC -3') sont mélangés à 50 µCi d'ATP <sup>32</sup>P, en présence de 10 unités de T4 polynucleotides kinase (Promega), dans un volume final de 25 µl d'eau. Les réactions sont ensuite incubées pendant 1 heure 30 minutes à 37°C, puis purifiées sur colonne spin 6 (Biorad) pour les oligonucléotides de taille inférieure à 20pb, et sur colonne Microbio spin P30 (Biorad) pour les oligonucléotides de taille supérieure à 20pb, en suivant les indications données par le fournisseur, et quantifiées sur gel d'agarose 1.2%.

### *b) Hybridation des oligonucléotides*

Les oligonucléotides marqués (10 pmol) PERV-U5A ou PERV-U5C sont alors hybridés avec 30 pmol de PERV-U5B (5'-AATGTAAGACCAGCAAAAAG -3'), alors que PERV-D est hybridé avec lui-même. La réaction d'hybridation est réalisée à 95°C pendant 5 minutes, puis à 37°C pendant 40 minutes, et sera enfin conservée à 4°C. L'hybridation des oligonucléotides est contrôlée sur un gel non dénaturant (30ml de 40% acrylamide: bisacrylamide (19:1), 20 ml TBE 5X, 500µl d'APS 10% et 80µl de TEMED, 50 ml H<sub>2</sub>O). Les oligonucléotides non hybridés sont également déposés sur le gel. La migration du gel se fait à 500V pendant 15 minutes dans une solution 10% acide acétique, 20% éthanol. Le gel est ensuite séché pendant 40 minutes à 80°C et exposé 1 h sur un film biomax AR (Kodak). Les oligonucléotides hybridés migrent plus haut que les oligonucléotides non hybridés.

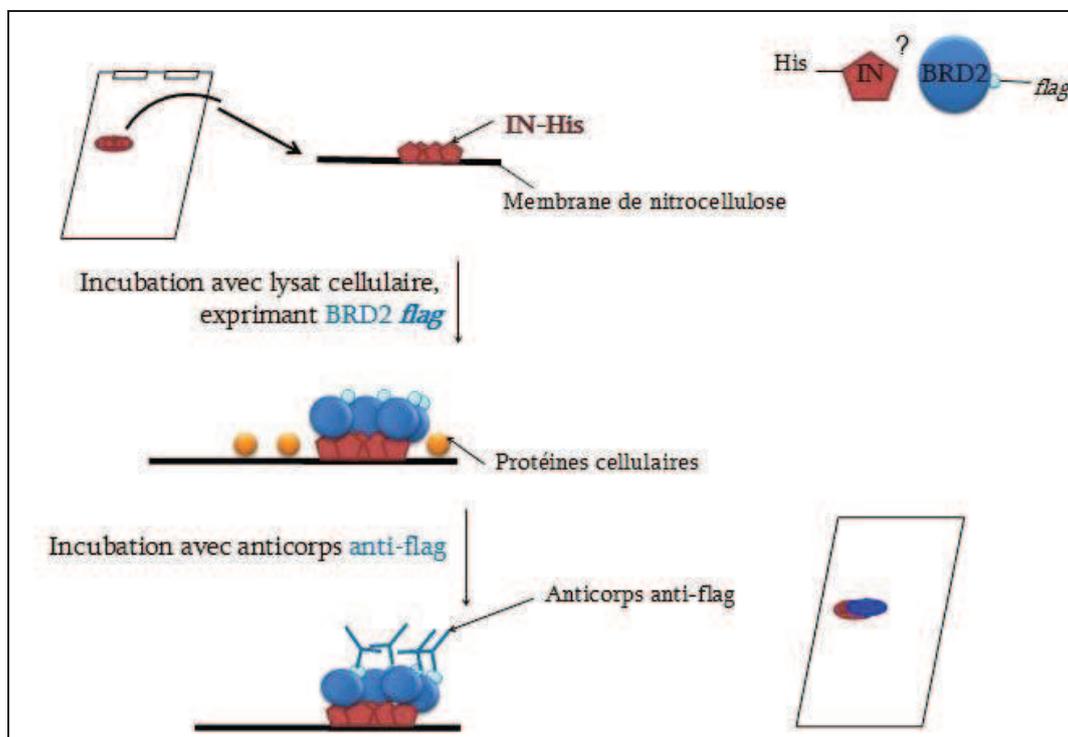
### *c) Réactions*

Pour les réactions de clivage en 3' et transfert de brins, 1 pmol d'oligonucléotides hybridés est incubé avec 100 ng d'intégrase. La réaction est effectuée pendant 1 heure 30 minutes à 37°C dans 10 µl d'un tampon contenant 20 mM MES, pH 6.2, 10 mM DTT, 10 mM MnCl<sub>2</sub>, 100 mM KCl, 10% (v/v) glycerol, pour les tests de clivage et de transfert de brins et dans 20 µl d'un tampon contenant 20 mM PIPES (pH 6.4), 1 mM DTT, 50 mM NaCl, 50 mM MnCl<sub>2</sub>, 7 mM CHAPS et 0.05% NP40, dans le test de désintégration. Les réactions sont

ensuite arrêtées par addition du tampon de dépôt 2X (10 ml formamide, 10 mg xylène cyanol, 200 µl EDTA 0.5 M à pH8, 10 mg bleu de bromophénol), suivie d'une incubation à 90°C pendant 5 minutes. Les produits issus de chaque réaction sont analysés par électrophorèse sur gel dénaturant contenant 15% polyacrylamide-7M urée. Après migration, le gel est séché et les produits des réactions sont observés par autoradiographie.

**VII- Test par Far Western Blot de l'interaction entre les INs (PERV ou ASLV) avec Brd2, entre les différents domaines de l'IN PERV et Brd2, et entre les INs (PERV ou ASLV) et le domaine C-terminal de Brd2, Brd3 et Brd4**

*A. Principe*



**Figure 20. Principe de Far western blot.** Séparation en électrophorèse sur gel SDS-PAGE dénaturant des protéines d'intérêt (représentées en rouge). Les protéines sont ensuite transférées par western blot sur une membrane de nitrocellulose, qui sera ensuite incubée avec la protéine partenaire (représentée en bleu) fusionnée à une étiquette pour laquelle existe un anticorps commercial. Si les deux protéines interagissent, suite à la détection de la protéine partenaire (en bleu), un signal sera obtenu au niveau de la taille de la protéine d'intérêt (en rouge).

Une expérience de far western blot typique débute par une séparation en électrophorèse sur gel SDS-PAGE dénaturant des protéines d'intérêt (x). Les protéines sont ensuite transférées par western blot sur une membrane de nitrocellulose ou PVDF. La membrane est alors bloquée puis incubée avec la protéine partenaire (y) fusionnée à une

étiquette pour laquelle existe un anticorps commercial conjugué à la HRP (horseradish peroxidase) ou l'AP (alkaline phosphatase). Si les deux protéines interagissent, suite à la détection de la protéine (y), un signal sera obtenu au niveau de la taille de la protéine (x) (figure 20).

Au contraire aux autres méthodes d'investigation d'interaction protéine-protéine, comme la co-IP (co-immunoprécipitation), ou FRET (fluorescent resonance energy transfer), la technique de far western blot possède l'avantage de pouvoir déterminer si deux protéines sont capables de se lier directement l'une à l'autre. Cette méthode possède également certains inconvénients comme la nécessité de renaturer la protéine sur son support afin de détecter avec succès l'interaction ; ceci sera discuté dans les étapes critiques (ci-dessous).

### *B. Les étapes critiques d'une analyse par far western blot*

#### *a) L'électrophorèse sur gel*

La migration des protéines sur SDS-PAGE sous conditions dénaturantes peut induire un changement de conformation de la protéine qui ne sera plus reconnaissable par la protéine partenaire, ou au contraire, reconnaitra la protéine d'une façon artificielle, conduisant à un résultat 'faux positif'. Dans ce dernier cas, la migration de la protéine sur gel devrait être réalisée sous conditions non-dénaturantes.

#### *b) Le transfert sur membrane*

Après séparation sur gel, les protéines sont transférées sur une membrane nitrocellulose ou PVDF. Or, les conditions de transfert peuvent altérer la conformation de la protéine, rendant le site d'interaction inaccessible. En général, la plupart des protéines subissent une renaturation après l'élimination du SDS. Dans les rares cas où la renaturation de la protéine n'a pas eu lieu, il sera donc nécessaire de réaliser une étape de dénaturation/renaturation (par du guanidinium hydrochloride) ou de détecter l'interaction protéine-protéine directement sur gel sans passer par l'étape de transfert.

#### *c) L'étape de blocage*

Suite au transfert des protéines sur la membrane, la procédure de western blot implique une étape d'incubation dans du tampon de blocage afin de bloquer les sites non occupés de la membrane. De plus, le tampon de blocage réduit les interactions non

spécifiques et participe à la renaturation de la protéine durant l'étape de détection. Un résultat optimal est généralement obtenu suite à un blocage de courte durée (environ 20 minutes), dans un tampon dont la concentration en lait ou BSA est élevée.

#### *d) Les conditions d'interaction et de lavage*

Les interactions protéines-protéines peuvent dépendre du pH, de la concentration en sels ou de la présence de certains co-facteurs durant l'incubation avec la protéine partenaire. Ces conditions doivent être maintenues durant les étapes de lavage pour préserver l'interaction jusqu'à la détection. Donc les lavages doivent être réalisés avec une agitation douce et sous une température de 4°C.

#### *e) Les méthodes de détection*

Il existe 4 méthodes pour la détection des interactions protéines-protéines durant un far western blot : - une détection directe de la protéine (x) par une protéine (y) radioactive.

- une détection indirecte par un anticorps dirigé contre la protéine (y).

- une détection indirecte par un anticorps dirigé contre une étiquette fusionnée à la protéine (y).

- une détection avec une protéine (y) biotinyllée et une enzyme (HRP/AP) marquée par l'avidine ou la streptavidine.

### *C. Protocole*

#### Matériel

- Le plasmide pCIneo3xFlag-Brd2 (préparé par k. Gallay (UMR754)) a été utilisé pour transfecter des cellules 293T ; chaque 2µg du plasmide sont utilisés pour transfecter x10 cellules 293T par CaCl<sub>2</sub>

- Les plasmides pGVB

- Production des domaines C-terminal de Brd2, Brd3 et Brd4 :

#### Méthode

- Pour vérifier les interactions entre les INs de PERV et d'ASLV et la protéine Brd2 entière, 1 à 2 µg d'intégrases pures sont déposées sur gel SDS-PAGE 10%. La migration sur gel est réalisée à 100 volts pendant 1h 30minutes. Le gel est ensuite transféré pendant 1 heure à 100 volts sur une membrane nitrocellulose. La membrane est ensuite bloquée pendant 20 minutes à 4°C dans du tampon contenant du PBST (1X Phosphate Buffered Saline Tween-20), ainsi que 10% de lait. La membrane est incubée pendant la nuit à 4°C avec 10 ml de PBST contenant 300µl de lysat cellulaire contenant la protéine Brd2 entière, ainsi qu'une tablette d'antiprotéases (complete EDTA-free, protease inhibitor cocktail) (Roche). La membrane est lavée 2 x 10 minutes dans du PBST et les protéines sont enfin révélées par une incubation avec 1/20000 anti-Flag M2 peroxydase (HRP) (Sigma). La révélation se fait sur un film (kodak) à l'aide du kit « Fast Western Blot Kits, SuperSignal West Pico » (Thermo scientific).

- Pour vérifier les interactions entre les différents domaines de l'IN PERV et Brd2, 20µl des lysats de cellules SF9 contenant les différents domaines de l'IN sont déposés sur gel SDS-PAGE 10% puis transférés sur une membrane nitrocellulose. La membrane est ensuite incubée avec 300µl de lysat cellulaire contenant la protéine Brd2 entière. La membrane est incubée avec 1/20000 anti-Flag M2 peroxydase (HRP).

- Pour vérifier les interactions entre les INs de PERV et d'ASLV et les domaines C-terminal de Brd2, Brd3 et Brd4, 2µg d'IN PERV et d'ASLV sont déposés sur gel SDS-PAGE puis transférés sur une membrane nitrocellulose qui est ensuite incubée avec 300µl de lysats bactériens contenant soit le domaine C-terminal de Brd2, soit le domaine C-terminal de Brd3, ou le domaine C-terminal de Brd4. La membrane est incubée avec 1/5000 d'anticorps anti-GST HRP (GE Healthcare).

## **RESULTATS**

## **I- Objectifs**

Pour permettre la vérification des interactions potentielles entre les INs et la protéine cellulaire Brd2, plusieurs conditions devaient être préalablement définies. Premièrement, les conditions d'activité de l'IN PERV devaient être déterminées et les techniques de production, de purification et des tests d'activités *in vitro* pour l'IN PERV devaient être développés. Dans ce cadre, nous souhaitions mettre au point la purification de l'IN de PERV à partir du système bactérien (comme cela se pratique couramment pour les INs actuellement), mais également tester la purification à partir du système baculovirus / cellules d'insecte, afin de voir si l'on pouvait obtenir des protéines plus actives suite aux modifications post-traductionnelles acquises par la production en cellules.

Les protéines ont été donc exprimées en bactérie et cellules SF9. Les protéines fusionnées à une étiquette 6xHistidine en N-ter ont été purifiées sur une colonne de chromatographie d'affinité HiTrap par FPLC (Fast Protein Liquid Chromatography). Les activités des protéines pures obtenues à partir des 2 systèmes de production différents ont été comparées par une suite de tests *in vitro* représentant les réactions de clivage et de transfert de brins réalisées par l'IN *in vivo*, et la réaction de désintégration qu'elle peut effectuer uniquement *in vitro*. Les résultats des tests d'activité de l'IN PERV seront détaillés dans l'article, alors que ceux de l'IN d'ASLV seront présentés dans la partie 'Données supplémentaires'.

Deuxièmement, une méthode de validation d'interaction protéine-protéine devait être mise au point afin de valider les interactions potentielles entre nos 2 enzymes virales actives et les protéines cellulaires Brd2 et TAB182.

Dans l'objectif de vérifier les interactions potentielles entre les INs et les 2 protéines cellulaires, différentes méthodes classiques de validation d'interaction entre protéines ont été tentées mais sans succès, comme le GST pull-down et la co-immunoprécipitation. Après plusieurs essais, ces tests ont été abandonnés faute de fixation non-spécifique de l'anticorps sur les billes, nous empêchant ainsi de vérifier les interactions.

Dans un second temps et suite à la purification des INs de PERV et d'ASLV, une expérience de «Far western blot» a été essayée. Le test consistait à faire migrer sur gel SDS-PAGE des protéines INs pures fusionnées à une étiquette 6xHis, et les transférer par western blot sur membrane nitrocellulose. Ensuite, il fallait incuber la membrane avec un lysat cellulaire contenant la protéine partenaire (Brd2 fusionnée à une étiquette 3xFlag). Cette méthode nous a permis de révéler une interaction spécifique entre l'IN PERV et Brd2. Enfin,

nous avons souhaité examiner de plus près cette interaction, en testant les différents domaines de l'IN et de Brd2. Pour cela, une délétion des différents domaines de l'IN de PERV a été réalisée, nous permettant ainsi de tester leur interaction avec Brd2 par « Far western Blot ».

## **II- Article**

Manuscrit en préparation

### **Characterization of the Porcine Endogenous Retrovirus (PERV) integrase activities *in vitro*, and validation of its interaction with the C-terminal domain of Bromodomain containing protein 2 (Brd2).**

Elsy Al ANDARY <sup>1-5</sup>, Kathy GALLAY <sup>2-5</sup>, Veronique BEVEN <sup>1</sup>, Karen MOREAU <sup>2-5</sup>, Anne-Cécile NIGNOL HERNANDEZ <sup>1</sup>, Yannick BLANCHARD <sup>1\*</sup> and Corinne RONFORT <sup>2-5,\*</sup>.

<sup>1</sup> Laboratoire de Génétique Virale et Biosécurité, AFSSA, BP53, 22440 Ploufragan, France

<sup>2</sup> Université de Lyon, F-69000 Lyon, France.

<sup>3</sup> Université Lyon 1, F-69007, Lyon, France.

<sup>4</sup> INRA, UMR754, Rétrovirus et Pathologie Comparée, F-69007; Lyon, France.

<sup>5</sup> UMS3444 BioSciences Gerland Lyon Sud, F-69007 Lyon, France.

**\* corresponding authors/**

Corinne Ronfort Tel : +33 437 287 629 Fax : +33 437 287 605 E-mail : [ronfort@univ-lyon1.fr](mailto:ronfort@univ-lyon1.fr).

Yannick Blanchard Tel : +33 296 016 297 Fax : +33 296 016 283 E-mail : [yannick.blanchard@anses.fr](mailto:yannick.blanchard@anses.fr).

## **Abstract**

Integrase (IN) has the key role in the integration of retroviral DNA into the host cell genome. Here, we described the purification of the Porcine Endogenous RetroVirus (PERV) IN obtained from two expression systems (bacteria and baculovirus), and analyzed their ability to mediate catalytic activities (3'-end processing, strand transfer and disintegration) *in vitro*, as well as to form dimers. Previously, we demonstrated that PERV shows a similar pattern of target site selection as Murine Leukemia Virus (MLV). In an attempt to identify host factors that might interact with PERV IN, we also investigated the interaction of Brd2 bromodomain containing protein 2 (Brd2), previously documented as partner of MLV IN, with PERV IN. Our results validate Brd2 as an interactant of PERV IN and point to the important role of the core/C-terminal domains of PERV IN in mediating the interaction with Brd2. In addition, no interaction was found with an Avian Sarcoma and Leukemia Virus (ASLV) IN.

## BACKGROUND

A critical step for retroviral replication is the stable integration of the provirus genome into the genome of its host. The viral enzyme integrase (IN), encoded by the *pol* gene, has a key role in this and carries out two reactions: (i) in the 3'-end processing (which proceeds in the cytoplasm), IN removes two nucleotides from the 3' termini of the viral long terminal repeats (LTRs) exposing a conserved CA-OH 3' ends with a two-bases 5'-overhang, and (ii) in strand transfer (which proceeds in the nucleus), IN joins the 3'ends of each recessed strand to the 5'end of the host DNA and results in a gapped intermediate (Engelman 2010). The integration process is completed by repair of the gaps flanking the integrated viral DNA and by elimination of the 5'-overhanging dinucleotides of the viral DNA ends. The repair of DNA gaps presumably involves host proteins (Smith and Daniel 2006; Espeseth, Fishel et al. 2011).

To study the integration event, *in vitro* assays have been developed by using purified IN enzyme and short oligonucleotides mimicking the viral LTR ends. *In vitro* studies showed that the IN enzyme is able to carry out both 3'-processing, strand transfer and disintegration (reverse of strand transfer) reactions. Biochemical and genetics studies showed that IN functions as a multimer which coordinates the integration of both viral DNA termini (Engelman, Bushman et al. 1993; van Gent, Vink et al. 1993) and most of recent data accumulate towards the IN functioning as a tetramer complex (Faure, Calmels et al. 2005; Li, Mizuuchi et al. 2006; Hare, Cherepanov et al. 2009; Hare, Di Nunzio et al. 2009; Michel, Crucifix et al. 2009; Hare, Gupta et al. 2010). The retroviral IN protein contains three functionally important domains: The N-terminal region (NTD) is characterized by an invariant HHCC (His-His/Cys-Cys) motif that binds zinc (Bushman, Engelman et al. 1993) which contributes to IN multimerization (Zheng, Jenkins et al. 1996). Several residues of the Human Immunodeficiency Virus type 1 (HIV-1) IN NTD were found to be involved in viral DNA binding (Heuer and Brown 1997; Heuer and Brown 1998; Zhao, McKee et al. 2008). The central core domain (CCD) is characterized by an active site consisting of a D-D(35)-E conserved motif. In addition to catalytic activities, the CCD is involved in specific recognition of viral ends (Jenkins, Esposito et al. 1997; Esposito and Craigie 1998; Gerton, Ohgi et al. 1998) and in binding host DNA (Katzman and Sudol 1995; Katzman and Sudol 1998; Katzman and Katz 1999). The C-terminal domain (CTD) is the less-conserved domain. It has non-specific DNA-binding activity (Mumm and Grandgenett 1991; Lutzke, Vink et al. 1994; Esposito and Craigie 1998) and it is also involved in the oligomerization of IN (Jenkins,

Engelman et al. 1996). An extended region of unknown function in the C-terminal domain has been identified in some IN proteins such as Murine Leukemia Virus (MLV) IN (Jonsson CB, Donzella G A et al. 1996) and Porcine Endogenous Retrovirus (PERV) IN.

A large number of HIV-1 IN-interacting host factors have been identified by analysing the components of the PIC or by yeast two hybrid screening (Kalpana, Marmon et al. 1994; Farnet and Bushman 1997; Lee and Craigie 1998; Parissi, Calmels et al. 2001; Rain JC, Cribier A et al. 2009). Human lens epithelium-derived growth factor (LEDGF/p75), the first host co-factor of HIV-1 IN whose role has been elucidated, was identified both in a yeast two-hybrid screen (Rain JC, Cribier A et al. 2009), and by its association with exogenously expressed HIV-1 IN in cells (Cherepanov, Maertens et al. 2003). Subsequent analysis of this factor has suggested an essential role for LEDGF/p75 in protecting IN from proteasomal degradation and for nuclear and chromosomal targeting of HIV-1 IN in human infected cells (Meehan and Poeschla 2010). This interaction may be responsible for preferential HIV integration within transcription units (Ciuffi, Llano et al. 2005).

The influence of host proteins on integration remain poorly defined for other retroviruses. Recently, a yeast two-hybrid screen using Mo-MLV IN as bait has been reported (Studamire B and Goff SP 2008). In this screen, the largest number of clones isolated corresponded to the cDNA for bromodomain containing protein 2 (Brd2/fsrg1/RING3) (Studamire B and Goff SP 2008). With Brd3, Brd4 and BrdT in mammals, Brd2 belongs to the BET family of proteins that are characterized by the presence of two bromodomains involved in recognition and binding of acetylated histones (Kanno, Kanno et al. 2004; LeRoy G, Rickards B et al. 2008). Brd2 was shown to be an essential factor as Brd2-deficient embryos die during mid-gestation development (Shang, Wang et al. 2009). Brd2 functions as a transcriptional co-activator and as a nuclear-localized kinase (Denis GV and Green MR 1996). It was also identified in a complex that contains, among others, E2F, histones, HDAC11, CBP, p300, Cyclin A2, and Swi/Snf chromatin remodeling complex member Brg-1 (Tamkun JW, Deuring R et al. 1992; Crowley TE, Kaine EM et al. 2002; Denis GV, McComb ME et al. 2006). Interestingly, Brd2 was also shown to interact with latency -associated nuclear antigen (LANA) Kaposi's sarcoma herpesvirus (KHSV) to tether the viral genome to mitotic chromosomes (Platt GM, Simpson GR et al. 1999; Mattsson K, Kiss C et al. 2002).

PERV are members of the gammaretrovirus genus. No pathology is formerly associated with PERV in its natural host but PERV has been shown to infect human cells *in*

*vitro* (Patience, Takeuchi et al. 1997). In previous studies (Moalic, Blanchard et al. 2006; Moalic, Felix et al. 2009), we characterized the integration profile of PERV in human cells. The integration profile of PERV showed an affinity for transcription start sites and CpG islands and displayed common patterns of target site preference with MLV and foamy virus (Mitchell RS, Beitzel BF et al. 2004; Nowrouzi A, Dittrich M et al. 2006; Trobridge GD, Miller DG et al. 2006). No studies of either PERV IN activities *in vitro* or of proteins interacting with PERV IN have been available until now. In the present study, we have expressed the PERV IN as a fusion protein with a 6xHistidine Tag in both *Escherichia coli* and insect SF9 cells. IN proteins from different viruses (HIV, Avian Sarcoma and Leukemia Viruses (ASLV), MLV) are usually purified from bacteria with high purity and activities. However, since eukaryotic systems are able to provide the proteins with many of the post-translational folding, processing and modification events required for their activities, we also produced PERV IN using the baculovirus expression system. After that, we analysed the catalytic activities of both PERV INs (3'-processing, strand transfer and disintegration) as well as their oligomeric state. Our result showed that zinc is required for dimerization of PERV IN and for the strand transfer activity of IN. Finally, in an attempt to identify host proteins that interact with PERV IN, we investigated the possibility of an interaction with Brd2. We used the far western blot method to validate this interaction. We also examined the interaction of Brd2 with the IN from Avian Sarcoma and Leukemia virus (ASLV), a virus which belongs to alpha-retrovirus genus.

## MATERIELS AND METHODS

### Construction of IN-expressing plasmids.

The full length PERV IN gene was amplified by PCR with Vent DNA polymerase (Biolabs), using the plasmid A14/220 (Oldmixon BA, Wood JC et al. 2002; Ericsson TA, Takeuchi Y et al. 2003) as template. All the PCR cycling conditions were as follows: 95 °C for 5 min, and then 30 cycles of 95 °C for 30 s, 55 °C for 60 s, and 72 °C for 1.5 min, and finally 72 °C for 7 min. In a first round of PCR, primers OGVB3074-F (CATATGATAGAAATGCCCAAAGCCCCAGAAC) and OGVB3074-R (CATATGTAACTTGAGGAGTTATTGACAGAG) were used. These primers were designed to include an NdeI and a HindIII restriction sites in the IN sequence. In a second round of PCR; primers OGVB3050-F (GCGGCCGCACATGGCACATCACCATCACCATCACATAGAAATGCCCAAAG) and OGVB3049-R (GAATTCTTAACTTGAGGAGTTATTGACAGAG) were used. These primers were designed to include a NotI and an EcoRI restriction sites in the IN sequence. After amplification, the PCR products were then digested with either NdeI and HindIII (set 1) or NotI and EcoRI (set 2) and products were inserted into the pET14b plasmid (Invitrogen) and into the PVL1392 plasmid (BD Baculogold), respectively. The resulting pGVB3045 construct (derivative of the pET14b plasmid) contains IN PERV sequence beyond an Histidine-Tag sequence and allows IN expression in bacteria. The resulting pGVB3030 construct (derivative of the PVL1392 plasmid) contains IN PERV sequence beyond an Histidine-Tag sequence and allows IN expression in SF9 insect cells.

Five PERV IN sub-domains (NTD, CCD, CTD, NTD + CCD and CCD + CTD) were also amplified by PCR using the same conditions as described above. The PCR products were then cloned into the PVL1392 plasmid (BD Baculogold) downstream of the polyhedrin promoter and beyond an Histidine-tag sequence, after digestion with EcoRI and NotI. The primers used to amplify the different sub-domains were as follow: For pGVB3040 (N-ter domain of PERV IN) we used the primers OGVB3050-F (GCGGCCGCACATGGCACATCACCATCACCATCACATAGAAATGCCCAAAG) and OGVB3071-R (GAATTCTAAGCATTAACCAGCTGGCAGGGC). For pGVB3041 (core domain of PERV IN) we used the primers OGVB3070-F (GCGGCCGCCATGGCACATCACCATCACCATCACCATCACAATCCTTCCAGAATG CCTCC) and OGVB3069-R (GAATTCTATACAGAAGCAATTTCTACCAATGG). For

pGVB3038 (C-ter domain of PERV) we used the primers OGVB3068-F (GCGGCCGCCATGGCACATCACCATCACCATCACCATCACGAAATTGCTTCTGTAC ATAGTGC) and OGVB3049-R (GAATTCTTAACTTGAGGAGTTATTGACAGAG). For pGVB3039 (N-ter/core domains of PERV IN) we used the primers OGVB3050-F (GCGGCCGCACATGGCACATCACCATCACCATCACATAGAAATGCCCAAAG) and OGVB3069-R (GAATTCTATACAGAAGCAATTTCTACCAATGG). For pGVB3037 (core/C-ter domains of PERV IN) we used the primers OGVB3070-F(GCGGCCGCCATGGCACATCACCATCACCATCACCATCACAATCCTTCCAGAATG CCTCC) and OGVB3049-R (GAATTCTTAACTTGAGGAGTTATTGACAGAG). The nucleotide sequence of each new construct was verified by DNA sequencing.

### **Construction of Brd2-expressing plasmids.**

The human Brd2 gene was amplified from 293T cells. Total RNAs were extracted by using the RNeasy kit (Qiagen) and then reverse transcribed with Reverse transcriptase of M-MLV (Promega). Brd2 sequence were amplified by two PCR using the Pfu Taq Pol (Promega): (i) by using primers 1\_fwd (ATGCTGCAAAACGTGACTCCCC) and 2\_rev (CGG TTC TCC ATC TTC CGC TTG) to amplify sequences from nucleotides 1 to 1217 (Genbank NM\_005104.3) (ii) by using primers 3-fwd (GGCCTGCATGACTACCATG) and 4- rev (TTAGCCTGAGTCTGAATCACTGGTGTG) to amplify sequences from nucleotides 1144 à 2406. The amplification reactions were performed by a 1 min at 95°C followed by 35 steps of 30 sec at 95°C, 30 sec at 62°C and 1 min 40 at 72°C and finished by a terminal step of 10 min at 72 °C. The PCR products were purified on Qiaquick column (Qiagen) and were then mixed, denaturated 5 min at 95°C, and let 30 minutes at laboratory temperature to allow hybridization. The mix was then used for a last elongation reaction of 15 min at 72°C. The resulting PCR product was used for a PCR adding Gateway cloning (att) sequences on each ends. Primers used were: 5\_fwd (GGGGACAACCTTTGTACAAAAAAGTTGGCATGCTGCAAAACGTGACTCCCC) and 6- rev (GGGGACAACCTTTGTACAAGAAAGTTGGTTAGCCTGAGTCTGAATC). Conditions of PCR were: one step of 1 min at 95°C followed by 35 steps of 30 sec at 95°C, 30 sec at 65°C and 2 min 40 sec at 72°C plus a final step of 10 min at 72°C. The 2406 bp long band corresponding to the full-length Brd2 sequence was further purified on a 0.8 % agarose gel with the GenElute gel extraction Kit (Sigma). Brd2 sequences were then cloned into the pDNOR223 plasmid (Invitrogen) using the Gateway technique (Invitrogen). Once the

sequence was checked, the Brd2 sequences were sub-cloned into the pCMVmycGW plasmid (Invitrogen).

The plasmid constructions of the C-terminal domains of Brd2, Brd3 and Brd4 proteins were generously provided by Dr. T.F. Schulz. Each domain was further sub-cloned in the pGEX-4T1 (Ottinger M, Christalla T et al. 2006) that contains GST (glutathione S-transferase). GST-Brd2/RING3 (C-terminus of Brd2/RING3 protein, aa 601-801) was cloned into the pGEX-4T1 vector using BamHI restriction site. GST-Brd3/ORF X (C-terminus of BRD3/ORF X protein, amino-acids 569-726) was cloned into the pGEX-4T1 vector using BamHI and EcoRI restriction sites. GST-Brd4/HUNK (C-terminus of Brd4/HUNK protein, amino-acids 607-722) was cloned into the pGEX-4T1 vector using BamHI and EcoRI restriction sites.

### **Purification of recombinant His-tagged PERV IN proteins from bacteria and insect cells.**

PERV IN<sub>bact</sub> was expressed in *E.coli* BL21 (DE3) bacteria (Novagen) transformed with the pGVB3045 plasmid. Bacteria were grown to an optical density at 600 nm of 0.6 in 1 liter of Luria broth at 37°C. Isopropyl-1-thio-b-D-galactopyranoside (IPTG) was then added to a final concentration of 1 mM and cells were grown for an additional 3 hours at 30°C and collected by centrifugation at 6.000 x g for 15 min. Bacterial cells were resuspended at 4°C in 10 ml of lysis buffer (20 mM Tris HCl, 500 mM NaCl, 10 mM Imidazole, 5 mM β-mercaptoethanol) for each 0.5 l of culture. After 30 min, one tablet of complete EDTA-free protease inhibitor cocktail (Roche) and 1.5 mg/ml of lysozyme were added to the lysate. Lysate was dounced until mixed then kept at 4°C for another 30 min. Samples were then disrupted by sonication. The lysates were clarified by centrifugation at 11.000 rpm for 30 min at 4°C. Solubilized extracts containing full-length PERV IN were filtered through 0.45 µm filters and were conserved at -20°C until purification (see below).

For production of PERV IN<sub>SF9</sub>, the pGVB3030 plasmid was co-transfected with the linearized BD BaculoGold DNA (BD Biosciences) into insect cells. *Spodoptera frugiperda* (SF9) insect cells seeded in a 25 cm<sup>2</sup> flask at 2 x 10<sup>6</sup> cells were transfected with a mixture of 0.5 µg bacmid and 15 µl Cellfectin reagent (Invitrogen). For high-titer protein expression, confluent SF9 cells were infected in a 250 ml volume at a cell density of 1.8 x 10<sup>6</sup> cells/ml. The supernatant was harvested 5 days post-transfection. The pellets of 1 x 10<sup>6</sup> cells were resuspended in 500 µl of insect cell lysis buffer (BD BaculoGold), in the presence of 1% (v/v)

of detergent Nonidet P-40 (NP40). Samples were then disrupted by sonication. The lysates were clarified by centrifugation at 13.000 rpm for 20 min at 4°C. Expression of proteins in infected cells was checked by western blot analysis using anti-histidine antibody (Pierce).

For purification of His-tagged PERV IN<sub>bact</sub> and PERV IN<sub>SF9</sub> proteins, the lysates supernatants were applied to Nickel charged-HiTrap IMAC HP column (GE Healthcare) equilibrated with buffer A (20 mM Tris HCl, 500 mM NaCl, 25 mM Imidazole, 5 mM β-mercaptoethanol). Unbound and weakly adsorbed proteins were washed from the affinity gel with 10 column volumes of buffer B (20 mM Tris HCl, 1 M NaCl, 25 mM Imidazole, 5 mM β-mercaptoethanol), then the His-tagged proteins were eluted with a gradient of buffer C (buffer B containing 1M imidazole) collecting 0.5 ml per fraction. Selected fractions were pooled and dialyzed against buffer C (20 mM Tris HCl, 1 M NaCl, 5 mM β-mercaptoethanol) 10% glycerol and 10 μM ZnCl<sub>2</sub>) and were stored at -80°C. Same conditions were used for PERV IN<sub>bact</sub> except that the dialysis buffer didn't contain ZnCl<sub>2</sub>.

### **Expression of the C-terminal domains of Brd2, Brd3 and Brd4 from bacteria.**

The C-terminal domains of Brd2, Brd3 and Brd4 proteins were expressed in bacteria. *E.coli* BL21 (DE3) (Novagen) were grown to an optical density at 600 nm of 0.8 in 1 liter of Luria broth at 37°C, and IPTG was added to a final concentration of 1mM. The bacteria were then grown for an additional 3h at 37°C and collected by centrifugation at 6000 x g for 15min. The pellets were stored at -80°C for later processing.

### **Protein-protein cross-linking.**

Cross-linking by Bis(sulfosuccinimidyl) suberate (Pierce) was performed as previously described (Moreau K, Faure C et al. 2003).

### **PERV IN activities.**

Oligonucleotides PERV-U5A (5'- CTTTTTGCTGGTCTTACATT -3') and PERV-U5Ac (5'- CTTTTTGCTGGTCTTACA -3') were used for 3'-processing and strand transfer reactions. Disintegration reactions were performed with an Y-oligomer substrate PERV-D (5'- TGTAAGATCTTACAGGCCCTTGGGCCGGCGCTTGCGCC -3') using PERV U5 viral DNA end-specific sequence in the viral portion. Oligonucleotides were 5'-end labelled with ATP <sup>32</sup>P by a T4 polynucleotide kinase (Invitrogen). PERV-U5A and PERV-U5Ac were then hybridized with PERV-U5B (5'- AATGTAAGACCAGCAAAAAG -3'). PERV-D was

hybridized with itself. Hybridization was performed as follow: oligonucleotides were denatured 5 min at 95°C, hybridized 40 min at 37°C and stored at 4°C. For 3'-processing and strand transfer reactions, 1 pmol of labelled substrate was mixed with 20 pmol of PERV IN in a reaction buffer containing 20 mM 4-morpholineethanesulfonic acid (MES, pH 6.2), 10 mM dithiothreitol (DTT), 10 mM MnCl<sub>2</sub>, 100 mM KCl, 10% (v/v) glycerol (Jonsson CB, Donzella GA et al. 1993; Yang F and Roth MJ 2001). Disintegration reactions were carried out with 1 pmol of PERV IN protein and 1 pmol of the disintegration substrate in a reaction buffer containing 20 mM 1,4-Piperazinediethanesulfonic acid (PIPES) (pH 6.4), 1 mM DTT, 50 mM NaCl, 50 mM MnCl<sub>2</sub>, 7 mM CHAPS, 0.05% NP40. The reaction mixture was incubated at 37°C for 1h and then stopped by adding 95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and heating at 90°C for 5 min. The reaction products were loaded on a 15% polyacrylamide-7M urea denaturing gel, followed by electrophoresis at 1.800 V. Reaction products were analysed by autoradiography.

### **Cell cultures and DNA transfection**

Lysates of human embryonic kidney 293T cells over-expressing the 3xFlag-Brd2 protein were used for Far Western analysis. Cells were transfected as previously described (Ronfort, De Breyne et al. 2004). 48 hours after transfection, cells were harvested and then lysed on ice for 20 min in 300 µl lysis buffer (20 mM Tris-HCl, pH 7.4, 180 mM NaCl, 1 mM EDTA, 0.5% NP40 and 1% protease inhibitors). Lysates were stored at -20 °C.

Lysates of insect SF9 cells expressing the full length IN PERV or its domains were used also for Far Western analysis. SF9 seeded in a 25 cm<sup>2</sup> flask at 2 x 10<sup>6</sup> cells were transfected with a mixture of 0.5 µg bacmid and 15 µl Cellfectin reagent (Invitrogen). The supernatant was harvested 5 days post-transfection.

### **Western and Far Western blotting analysis.**

Proteins were suspended in an equal volume of laemmli buffer (60 mM Tris-Cl pH 6.8, 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.01% bromophenol blue) and were heated for 5 min at 95°C. The proteins were then electrophoresed in SDS-denaturing polyacrylamide gels and were electrically transferred to nitrocellulose membranes (100 Volts, 1 hour). For Western analysis, the blots were blocked for 1 hour in Tris-buffered saline containing 0.05% Tween 20 (TBS-T) + 25mg/ml BSA (Bovine Serum Albumin), and then incubated with either an His-probe-HRP antibody (working dilution 1/5000) (Thermo Scientific) or with the Anti-

Flag M2 peroxylase (HRP) antibody (working dilution 1/20000)(Sigma), or with the Anti-GST HRP conjugate (working dilution 1/5000)(GE Healthcare). The western Blot detection ECL kit (Thermo scientific) was used for revelation.

For Far Western blot analysis, 1  $\mu$ g of purified IN proteins were subjected to 10% SDS-PAGE and transferred to nitrocellulose membrane. The blot was blocked with PBST (Phosphate-buffered saline containing 0.05% Tween 20) containing 10% skim milk for 20 min and incubated overnight at 4°C with 300  $\mu$ l of 293T lysates containing the overexpressed Brd2 protein in 10 ml of PBST supplemented with 1 tablet of complete EDTA-free, protease inhibitor cocktail (Roche). The membrane was washed by gentle agitation two times for 10 min each with PBST to remove unbound proteins, and was finally subjected to a Western analysis using the anti-Flag-M2 peroxidase antibody or anti-GST HRP conjugate antibody (described above).

## RESULTS AND DISCUSSION

### Expression and purification of WT PERV IN proteins

PERV IN coding-sequences were subcloned into pET14b and PVL1392 plasmids and expressed as an N-terminal hexa-histidine tagged protein in *E. coli* and SF9 insect cells, respectively. Maximal IN expression was observed at 3 hours post-induction in *E. coli* and 48 hours post infection in SF9 cells (data not shown).

To purify PERV IN from *E. Coli*, we first used our protocol of ASLV IN purification as previously described (Moreau, Faure et al. 2003; Moreau, Faure et al. 2004; Moreau, Charmetant et al. 2009; Ballandras, Moreau et al. 2011). This initial attempt was unsuccessful. PERV IN was insoluble maybe due to the difference in amino acid composition among IN proteins. Therefore, purification was performed according to protocols previously developed for MLV (Jonsson CB, Donzella GA et al. 1993; Yang F, Leon O et al. 1999). The use of the optimal conditions defined for MLV generated high yields (over 500 mg per liter of culture) of pure PERV IN proteins (named PERV IN<sub>bact</sub>) suitable for biochemical analysis.

For initial protein extraction from SF9, effective lysis strategies were developed using non-ionic and zwitterionic detergents. After that, purification was performed using the same conditions described for PERV IN<sub>bact</sub>.

Therefore, both proteins (IN<sub>bact</sub> and IN<sub>SF9</sub> purified from *E. Coli* and insect cells, respectively) were purified by a one-step nickel-chelating chromatography. After extensive washing of the columns, IN was eluted with a linear imidazole gradient (25 mM-1M) giving yields exceeding 2 mg of purified protein, per litre of culture for *E. coli*. Fewer yields of purified proteins were obtained from SF9 expression system because of the limited solubility (the protein was found at 50% in the pellet of cell lysate according to results obtained from western blot analysis (data not shown)). The addition of detergent (Igepal CA-630, Sigma) was therefore an absolute requirement for protein extraction from SF9 cells. After purification, both proteins (IN<sub>bact\_zc</sub> and IN<sub>SF9</sub>) were dialyzed for 18 hours in the presence of 10  $\mu$ M zinc required for the efficiency of the IN activity (Yang F and Roth MJ 2001), whereas IN<sub>bact</sub> was dialyzed in the absence of zinc.

Our results showed that IN proteins expressed in bacteria were produced at much higher level than the protein expressed in baculovirus system. Proteins were then deposited on

a 12% SDS polyacrylamide gels (**Fig.1A**). IN<sub>bact</sub> and IN<sub>SF9</sub> were routinely greater than 90% pure.

### **Zinc promotes dimerization of PERV IN<sub>bact</sub>**

IN<sub>bact\_zc</sub> protein was tested for its ability to form a complex by protein–protein cross-linking. IN<sub>bact\_zc</sub> was incubated with the Bis (sulfosuccinimidyl) suberate (BS<sup>3</sup>) cross-linker. Reaction products were separated by SDS-PAGE and revealed by Western blot using an anti-Histidine antibody (**Fig. 1B**). In the absence of BS<sup>3</sup>, only a single band with apparent molecular weight of 49 Kda corresponding to the monomeric form of PERV IN was observed (lane 3). In the presence of BS<sup>3</sup>, the expected molecular weights of IN monomers and dimers were observed (lane 4). This dimeric state was zinc-dependent. Indeed, in the absence of 10 μM zinc in the dialysis solution, PERV IN yielded no detectable dimers (lane 2). Our experiment revealed a dimer form of the protein strongly dependent on the presence of the divalent cation zinc.

### ***In vitro* activities catalyzed by PERV IN<sub>bact</sub> and PERV IN<sub>SF9</sub>**

The IN PERV INs were then assayed for 3' processing, strand transfer and disintegration *in vitro* activities (**Fig. 2**). These assays rely on the use 5' radio-labelled small-size oligonucleotide substrates.

The 3'-processing and strand transfer activities of PERV IN were assayed under the conditions described in Methods, which were based on the optimal conditions defined for MLV IN by Roth and al. (Jonsson CB, Donzella GA et al. 1993; Yang F and Roth MJ 2001). In these assays, 1 pmol of a 20 mer 5'-labeled blunt-ended oligonucleotide mimicking the terminal U5 of the PERV LTR was used as the substrate for the reaction. Reactions were performed in 100 mM KCl, at pH 6.2 and in the presence of Mn<sup>2+</sup> as a metallic cofactor. As shown in **Fig. 2A** (lane 1), the three proteins yielded a product running exactly coincident with the expected -2 nt processed product (lanes 2, 3 and 4) while such a product was not detected in absence of IN (lane 1). In this test, PERV IN<sub>bact</sub> and PERV IN<sub>bact\_zc</sub> were reproducibly the most efficient. A slightly fainter signal was detected in this assay for PERV IN<sub>SF9</sub>. Proteins were also tested for strand transfer activity using a 5'-labelled pre-cleaved oligonucleotide substrate (**Fig. 3B**). Using the optimal conditions defined for MLV IN by Roth and al. (Jonsson CB, Donzella GA et al. 1993), a part of the pre-cleaved strand transfer substrate was converted to integration product by the PERV IN<sub>SF9</sub> (lane 2) PERV IN<sub>bact\_zc</sub>

(lane 3) and, by PERV IN<sub>bact</sub> but to a lesser extent (lane 4) while no such products were detected in absence of IN (lane 1).

In the disintegration reaction (**Fig. 3C**), we used a Y-oligomer substrate that mimics the product of a strand transfer reaction with oligonucleotides corresponding to PERV DNA end sequences. Disintegration is the reverse of the strand transfer reaction and results in two products, mimicking the labelled viral DNA end and the unlabeled target DNA. The conditions of the reaction were similar to those used for 3'-end processing and strand transfer, except that the reaction was performed at pH 6.4 in 50 mM NaCl. In absence of IN, no disintegration products are observed (lane 1). The reaction in the presence of PERV IN<sub>SF9</sub> (lane 2), PERV IN<sub>bact zc</sub> (lane 3) and PERV IN<sub>bact</sub> (lane 4) yielded to the expected 14-long labelled DNA product. As for the 3'-processing reaction, PERV IN<sub>bact</sub> was more efficient than PERV IN<sub>SF9</sub> and PERV IN<sub>bact zc</sub>.

Altogether, these data show PERV IN<sub>bact zc</sub> activity is higher than that of PERV IN<sub>SF9</sub>, as purified in our experimental conditions. Both have higher integration activities than the PERV IN purified from insect cells. The poorly efficient production of PERV IN from insect cells requires the addition of detergent. Knowing that the use of non-ionic detergent might affect the folding of the protein, it could explain the slightly lower *in vitro* 3'-processing and strand transfer activities obtained with PERV IN expressed in SF9 cells compared to PERV IN purified in bacteria in presence of zinc.

### **Interaction *in vitro* of human BRD2 protein with PERV IN but not with ALSV IN in far western Blot.**

Brd2 has been already shown to interact with the MLV IN. In a two-hybrid assay, nine isolates of the yeast two-hybrid screen using MLV IN as bait, corresponded to the cDNA for Brd2 (Studamire B and Goff SP 2008). We thus questioned if Brd2 would be also a partner of PERV IN (another gamma-retrovirus). We also investigated the interaction of Brd2 with IN from an avian alpha-retrovirus (Rous Associated Virus type 1 (RAV-1) we have previously purified (Moreau, Faure et al. 2002; Moreau, Faure et al. 2004). For this purpose, we first tried to use the GST pull down and co-immuno-precipitation methods. These attempts were unsuccessfully due to non-specific-binding of IN proteins to the beads. Therefore, we managed to analyse *in vitro* the potential interaction between Brd2 and both INs by using the Far western blot method. PERV IN<sub>bact</sub> and Rous Associated Virus type 1 IN protein purified

from bacteria (Moreau, Faure et al. 2003; Moreau, Faure et al. 2004; Moreau, Charmetant et al. 2009) were subjected to migration on SDS-PAGE and immobilized on a nitrocellulose membrane. A first membrane was revealed with an anti-His antibody to reveal the 49 kDa PERV IN<sub>bact</sub> and the 37 kDa RAV1 IN proteins (ASLV IN<sub>bact</sub>) (**Fig. 3B**, lanes 1 to 2). A lysate of 293T transiently expressing a flag-Brd2 (48 hours) was used to probe INs on a second membrane (lane 3). Brd2 retention on membrane was then revealed with an anti-Flag antibody (**Fig. 3B**, lanes 4 to 6). A flag-signal was revealed at position corresponding to the PERV IN<sub>bact</sub> (**Fig. 3B**, lane 4), confirming the interaction between PERV IN<sub>bact</sub> and Brd2. By contrast, no flag-signal was revealed with RAV-1 IN (lane 5) indicating that RAV-1 IN did not interact with human Brd2. A positif control corresponding to the 3xFlag-Brd2 protein was also revealed on the membrane (lane 6).

Therefore, Brd2 is a specific partner to at least two genera of retroviruses: gamma-retroviruses (MLV (Studamire B and Goff SP 2008) and PERV INs (herein)) and lentiviruses (HIV-IN (Studamire B and Goff SP 2008)). We show here that human Brd2 is not a partner of ASLV IN, at least in vitro. As it is known that ASLV retroviral vector are able to infect human cells and to efficiently integrate into human DNA (Mitchell, Beitzel et al. 2004; Torne-Celer, Moreau et al. 2008) Brd2 is probably not a co-factor of alpha-retrovirus IN, at least in mammals cells.

### **Delineation of region(s) of PERV IN required for its interaction with Brd2**

Furthermore, we decided to determine the IN domain(s) of PERV IN responsible for its interaction with Brd2. In this experiment, truncated PERV IN mutants were generated. Therefore, five IN truncated mutants were constructed that included IN NTD (aa1-101), IN CCD (aa102-181), IN CTD (aa182-418), IN NTD/CTD (aa1-181) and IN CCD/CTD (aa102-418) (**Fig. 3C**). The 6xhis-fusion proteins were expressed in SF9 cells. Lysates of SF9 cells were deposited on SDS-PAGE gel and transferred to a nitrocellulose membrane. INs were revealed with an anti-His antibody on the first membrane. The proteins were detected at the expected molecular weights ranging from 15 to 35 kDa (**Fig.3D**, lanes 4 to 5). The second membrane was incubated with 293T cells lysates expressing the flag-Brd2 protein. The interaction was revealed by using a Flag antibody. Interestingly, the far western analysis revealed that only the IN CCD/CTD protein was able to bind Brd2 (**Fig.3D**, lane 1 right). By contrast, neither the CTD (lane 2), the CCD (lane 5), NTD (lane 4) nor a truncated two-domains protein (NTD/CCD) show any interaction with Brd2.

Our results highlighted the specific role of the CCD/CTD domains of PERV IN in binding the cellular Brd2. The deletion of one of these two domains compromised the association with Brd2; this is probably due to a conformational change in the protein structure. Alternatively, Brd2 need both the CCD/CTD domains to establish a stable association.

### **Mapping for Brd2 interacting domain by far western blot analysis**

Previous studies have indicated that the C-terminal domain of Brd2 is involved in protein-protein interactions (Florence B and Faller DV 2001). Furthermore; this domain has previously been shown to interact with MLV IN (Studamire B and Goff SP 2008). Therefore, we analysed the interaction between INs and the extra-terminal (ET) C-terminal of Brd2. The study was further extended to other proteins of the BET family. Thus, we also tested Brd3 and Brd4 C-terminal domains (**Fig. 3G**). All the GST fusion proteins were produced in E.coli (**Fig. 3F**, lanes 1 to 4). For the Far western Blot, PERV IN<sub>bact</sub> and ASLV IN proteins were separated by SDS-PAGE and transferred to nitrocellulose membrane and revealed with an anti-his antibody (**Fig. 3G**, lanes 1 and 2). The membrane was subsequently incubated with bacterial lysate containing either the GST flag alone, or GST- Brd2 C-terminal protein, or the C-terminal domain of Brd3 and Brd4. An interaction was detected by using the anti-GST antibody. The far western analysis did not show any interaction between the GST fragment and the 6xHis-INs on gel. The far western analysis showed that PERV IN<sub>bact</sub> was able to interact with the C-ter domain of Brd2 (**Fig.3G**, lane 4). Whereas, no interaction was found between PERV IN<sub>bact</sub> and the C-ter domains of Brd3 or Brd4 (results not shown). This analysis also confirmed the absence of interaction of ASLV IN with the C-terminal domain of Brd2 (lane 5).

## CONCLUSION

We report herein the interaction of the CCD/CTD portion of PERV IN with the Extra-terminal domain of the cellular Brd2 protein but not with ET domain of other proteins of the BET family (Brd3 and brd4). Other have reported the interaction of the Brd2 ET domain with IN of MLV (Studamire B and Goff SP 2008). Interestingly, Brd4, by its interaction with virus-encoded E2 protein is used as a cellular adaptor by human papillomaviruses (HPV) for anchoring viral genomes to mitotic chromosomes (Jang MK, Kwon D et al. 2009). Should a similar role be assigned to BRD2 for anchoring gammaretroviral genome to the mitotic chromosomes? Altogether, these observations suggest that the chromatin-binding Brd2 protein may play a role in retroviral integration. Finally, this study can be considered as a first step in the understanding of the precise interaction between cellular proteins and PERV IN.

**ACKNOWLEDGMENTS** We thank Dr. T.F. Schulz (Institut für Virologie, Medizinische Hochschule Hannover, Carl-Neuberg Str. 1, 30625 Hannover, Germany) for providing us the plasmids pGST-4T1 containing GST, pGST-Brd2/RING3, pGST-Brd3/ORF X and pGST-Brd4/HUNK. We are grateful to the “Production and Analysis of Proteins” platform of UMS3444 BioSciences Gerland Lyon Sud. EAE was supported by a doctorate fellowship from INRA and ANSES. We also thank Lise-Marie Donnio for technical assistance.

## REFERENCES

- Ballandras A, Moreau K, Robert X, Confort M, Merceron R, Haser R, Ronfort C, 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(8):e23032.
- Bushman FD, Engelman A, Palmer I, Wingfield P, Craigie R. 1993. Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. *Proceedings of the National Academy of Sciences of the United States of America* 90(8):3428-3432.
- Cherepanov P, Maertens G, Proost P, Devreese B, Van Beeumen J, Engelborghs Y, De Clercq E, Debysers Z. 2003. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *The Journal of biological chemistry* 278(1):372-381.
- Ciuffi A, Llano M, Poeschla E, Hoffmann C, Leipzig J, Shinn P, Ecker JR, Bushman F. 2005. A role for LEDGF/p75 in targeting HIV DNA integration. *Nat Med* 11(12):1287-1289.
- Crowley TE, Kaine EM, Yoshida M, Nandi A, Wolgemuth DJ. 2002. Reproductive cycle regulation of nuclear import, euchromatic localization, and association with components of Pol II mediator of a mammalian double-bromodomain protein. *Mol Endocrinol* 16(8):1727-1737.
- Denis GV, Green MR. 1996. A novel, mitogen-activated nuclear kinase is related to a *Drosophila* developmental regulator. *Genes Dev* 10(3):261-271.
- Denis GV, McComb ME, Faller DV, Sinha A, Romesser PB, Costello CE. 2006. Identification of transcription complexes that contain the double bromodomain protein Brd2 and chromatin remodeling machines. *J Proteome Res* 5(3):502-511.
- Engelman A. 2010. Reverse transcription and integration. In: Kurth R, Bannert N, editors. *Retroviruses: Molecular biology, Genomics and Pathogenesis*. Norfolk, UK: Caister Academic Press. p 129-159.
- Engelman A, Bushman FD, Craigie R. 1993. Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. *The EMBO journal* 12(8):3269-3275.
- Ericsson TA, Takeuchi Y, Templin C, Quinn G, Farhadian SF, Wood JC, Oldmixon BA, Suling KM, Ishii JK, Kitagawa Y, Miyazawa T, Salomon DR, Weiss RA, Patience C. 2003. Identification of receptors for pig endogenous retrovirus. *Proc Natl Acad Sci U S A* 100(11):6759-6764.
- Espeseth AS, Fishel R, Hazuda D, Huang Q, Xu M, Yoder K, Zhou H. 2011. siRNA screening of a targeted library of DNA repair factors in HIV infection reveals a role for base excision repair in HIV integration. *PLoS ONE* 6(3):e17612.
- Esposito D, Craigie R. 1998. Sequence specificity of viral end DNA binding by HIV-1 integrase reveals critical regions for protein-DNA interaction. *The EMBO journal* 17(19):5832-5843.
- Farnet CM, Bushman FD. 1997. HIV-1 cDNA integration: requirement of HMG I(Y) protein for function of preintegration complexes in vitro. *Cell* 88(4):483-492.
- Faure A, Calmels C, Desjobert C, Castroviejo M, Caumont-Sarcos A, Tarrago-Litvak L, Litvak S, Parissi V. 2005. HIV-1 integrase crosslinked oligomers are active in vitro. *Nucleic acids research* 33(3):977-986.
- Florence B, Faller DV. 2001. You bet-cha: a novel family of transcriptional regulators. *Front Biosci* 6:1008-1018.

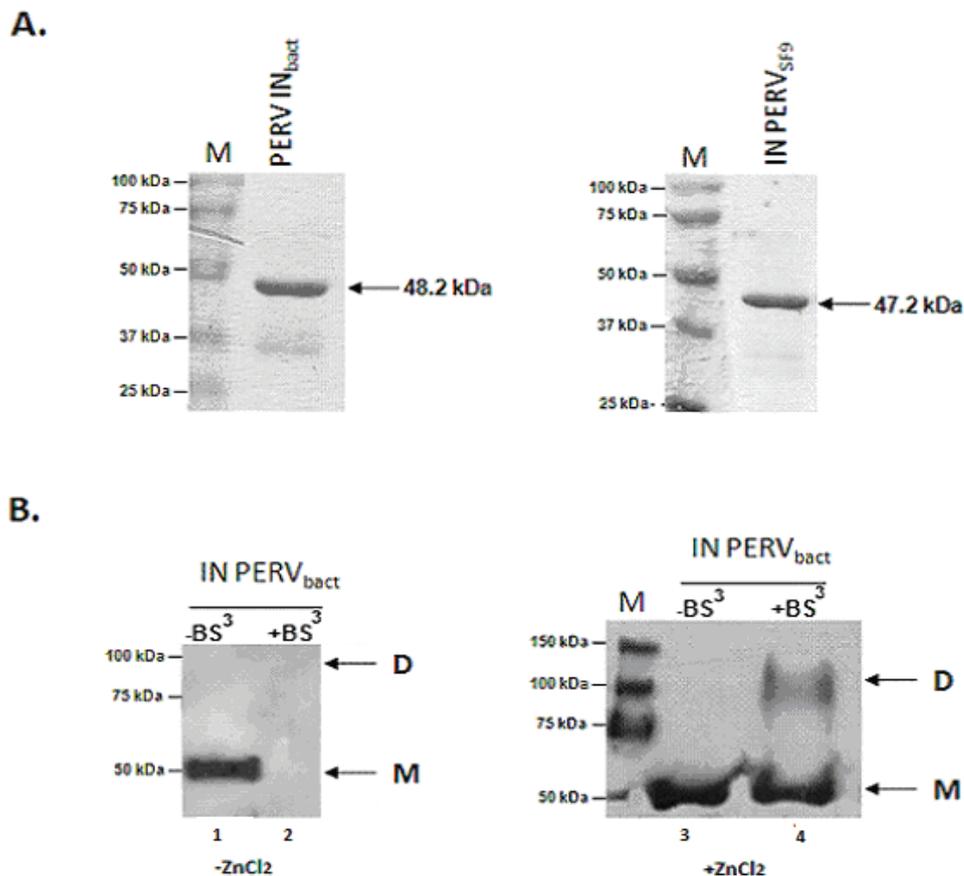
- Gerton JL, Ohgi S, Olsen M, DeRisi J, Brown PO. 1998. Effects of mutations in residues near the active site of human immunodeficiency virus type 1 integrase on specific enzyme-substrate interactions. *Journal of virology* 72(6):5046-5055.
- Hare S, Cherepanov P, Wang J. 2009a. Application of general formulas for the correction of a lattice-translocation defect in crystals of a lentiviral integrase in complex with LEDGF. *Acta crystallographica* 65(Pt 9):966-973.
- Hare S, Di Nunzio F, Labeja A, Wang J, Engelman A, Cherepanov P. 2009b. Structural basis for functional tetramerization of lentiviral integrase. *PLoS pathogens* 5(7):e1000515.
- Hare S, Gupta SS, Valkov E, Engelman A, Cherepanov P. 2010. Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* 464(7289):232-236.
- Heuer TS, Brown PO. 1997. Mapping features of HIV-1 integrase near selected sites on viral and target DNA molecules in an active enzyme-DNA complex by photo-cross-linking. *Biochemistry* 36(35):10655-10665.
- Heuer TS, Brown PO. 1998. Photo-cross-linking studies suggest a model for the architecture of an active human immunodeficiency virus type 1 integrase-DNA complex. *Biochemistry* 37(19):6667-6678.
- Jang MK, Kwon D, McBride AA. 2009. Papillomavirus E2 proteins and the host BRD4 protein associate with transcriptionally active cellular chromatin. *J Virol* 83(6):2592-2600.
- Jenkins TM, Engelman A, Ghirlando R, Craigie R. 1996. A soluble active mutant of HIV-1 integrase: involvement of both the core and carboxyl-terminal domains in multimerization. *The Journal of biological chemistry* 271(13):7712-7718.
- Jenkins TM, Esposito D, Engelman A, Craigie R. 1997. Critical contacts between HIV-1 integrase and viral DNA identified by structure-based analysis and photo-crosslinking. *The EMBO journal* 16(22):6849-6859.
- Jonsson CB, Donzella G A, Gaucan E, Smith C M, Roth MJ. 1996. Functional domains of Moloney murine leukemia virus integrase defined by mutation and complementation analysis. *J Virol* 70:4585-4597.
- Jonsson CB, Donzella GA, Roth MJ. 1993. Characterization of the forward and reverse integration reactions of the Moloney murine leukemia virus integrase protein purified from *Escherichia coli*. *J Biol Chem* 268(2):1462-1469.
- Kalpana GV, Marmon S, Wang W, Crabtree GR, Goff SP. 1994. Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. *Science* (New York, NY 266(5193):2002-2006.
- Kanno T, Kanno Y, Siegel RM, Jang MK, Lenardo MJ, Ozato K. 2004. Selective recognition of acetylated histones by bromodomain proteins visualized in living cells. *Mol Cell* 13(1):33-43.
- Katzman M, Katz RA. 1999. Substrate recognition by retroviral integrases. *Adv Virus Res* 52:371-395.
- Katzman M, Sudol M. 1995. Mapping domains of retroviral integrase responsible for viral DNA specificity and target site selection by analysis of chimeras between human immunodeficiency virus type 1 and visna virus integrases. *Journal of virology* 69(9):5687-5696.
- Katzman M, Sudol M. 1998. Mapping viral DNA specificity to the central region of integrase by using functional human immunodeficiency virus type 1/visna virus chimeric proteins. *Journal of virology* 72(3):1744-1753.
- Lee MS, Craigie R. 1998. A previously unidentified host protein protects retroviral DNA from autointegration. *Proceedings of the National Academy of Sciences of the United States of America* 95(4):1528-1533.

- LeRoy G, Rickards B, Flint SJ. 2008. The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. *Mol Cell Biol* 30(1):51-60.
- Li M, Mizuuchi M, Burke TR, Jr., Craigie R. 2006. Retroviral DNA integration: reaction pathway and critical intermediates. *The EMBO journal* 25(6):1295-1304.
- Lutzke RA, Vink C, Plasterk RH. 1994. Characterization of the minimal DNA-binding domain of the HIV integrase protein. *Nucleic acids research* 22(20):4125-4131.
- Mattsson K, Kiss C, Platt GM, Simpson GR, Kashuba E, Klein G, Schulz TF, Szekely L. 2002. Latent nuclear antigen of Kaposi's sarcoma herpesvirus/human herpesvirus-8 induces and relocates RING3 to nuclear heterochromatin regions. *J Gen Virol* 83(1):179-188.
- Meehan AM, Poeschla EM. 2010. Chromatin tethering and retroviral integration: recent discoveries and parallels with DNA viruses. *Biochimica et biophysica acta* 1799(3-4):182-191.
- Michel F, Crucifix C, Granger F, Eiler S, Mouscadet JF, Korolev S, Agapkina J, Ziganshin R, Gottikh M, Nazabal A, Emiliani S, Benarous R, Moras D, Schultz P, Ruff M. 2009. Structural basis for HIV-1 DNA integration in the human genome, role of the LEDGF/P75 cofactor. *The EMBO journal* 28(7):980-991.
- Mitchell RS, Beitzel BF, Schroder AR, Shinn P, Chen H, Berry CC, Ecker JR, Bushman FD. 2004. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol* 2(8):1127-1137.
- Mitchell RS, Beitzel BF, Schroder AR, Shinn P, Chen H, Berry CC, Ecker JR, Bushman FD. 2004. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol* 2(8):E234.
- Moalic Y, Blanchard Y, Felix H, Jestin A. 2006. Porcine endogenous retrovirus integration sites in the human genome: features in common with those of murine leukemia virus. *Journal of virology* 80(22):10980-10988.
- Moalic Y, Felix H, Takeuchi Y, Jestin A, 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. *Journal of virology* 83(4):1920-1929.
- Moreau K, Faure C, Violot S, Verdier G, 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(22):4426-4438.
- Moreau K, Charmetant J, Gallay K, Faure C, Verdier G, Ronfort C. 2009. Avian sarcoma and leukemia virus (ASLV) integration in vitro: mutation or deletion of integrase (IN) recognition sequences does not prevent but only reduces the efficiency and accuracy of DNA integration. *Virology* 392(1):94-102.
- Moreau K, Faure C, Verdier G, 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(9):1761-1778.
- Moreau K, Faure C, Violot S, Gouet P, Verdier G, 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(2):566-581.
- Moreau K, Faure C, Violot S, Verdier G, 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. *European journal of biochemistry / FEBS* 270(22):4426-4438.
- Mumm SR, Grandgenett DP. 1991. Defining nucleic acid-binding properties of avian retrovirus integrase by deletion analysis. *Journal of virology* 65(3):1160-1167.

- Nowrouzi A, Dittrich M, Klanke C, Heinkelein M, Rammling M, Dandekar T, von Kalle C, Rethwilm A. 2006. Genome-wide mapping of foamy virus vector integrations into a human cell line. *J Gen Virol* 87:1339-1347.
- Oldmixon BA, Wood JC, Ericsson TA, Wilson CA, White-Scharf ME, Andersson G, Greenstein JL, Schuurman HJ, Patience C. 2002. Porcine endogenous retrovirus transmission characteristics of an inbred herd of miniature swine. *J Virol* 76(6):3045-3048.
- Ottinger M, Christalla T, Nathan K, Brinkmann MM, Viejo-Borbolla A, Schulz TF. 2006. Kaposi's 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. *J Virol* 80(21):10772-10786.
- Parissi V, Calmels C, De Soultrait VR, Caumont A, Fournier M, Chaignepain S, Litvak S. 2001. Functional interactions of human immunodeficiency virus type 1 integrase with human and yeast HSP60. *Journal of virology* 75(23):11344-11353.
- Patience C, Takeuchi Y, Weiss RA. 1997. Infection of human cells by an endogenous retrovirus of pigs. *Nat Med* 3(3):282-286.
- Platt GM, Simpson GR, Mittnacht S, Schulz TF. 1999. Latent nuclear antigen of Kaposi's sarcoma-associated herpesvirus interacts with RING3, a homolog of the *Drosophila* female sterile homeotic (*fsh*) gene. *J Virol* 73(12):9789-9795.
- Rain JC, Cribier A, Gérard A, Emiliani S, Benarous R. 2009. Yeast two-hybrid detection of integrase-host factor interactions. *Methods* 47(4):291-297.
- Ronfort C, De Breyne S, Sandrin V, Darlix JL, Ohlmann T. 2004. Characterization of two distinct RNA domains that regulate translation of the *Drosophila* gypsy retroelement. *Rna* 10(3):504-515.
- Shang E, Wang X, Wen D, Greenberg DA, Wolgemuth DJ. 2009. Double bromodomain-containing gene *Brd2* is essential for embryonic development in mouse. *Dev Dyn* 238(4):908-917.
- Smith JA, Daniel R. 2006. Following the path of the virus: the exploitation of host DNA repair mechanisms by retroviruses. *ACS chemical biology* 1(4):217-226.
- Studamire B, Goff SP. 2008. Host proteins interacting with the Moloney murine leukemia virus integrase: multiple transcriptional regulators and chromatin binding factors. *Retrovirology* 5(48).
- Tamkun JW, Deuring R, Scott MP, Kissinger M, Pattatucci AM, Kaufman TC, Kennison JA. 1992. *brhma*: a regulator of *Drosophila* homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell Mol Life Sci* 68(3):561-572.
- Torne-Celer C, Moreau K, Faure C, Verdier G, Ronfort C. 2008. Gene transfer into mammalian cells using a self-deleting avian leukemia and sarcoma virus-based vector. *Intervirology* 51(6):447-452.
- Trobridge GD, Miller DG, Jacobs MA, Allen JM, Kiem HP, Kaul R, Russell DW. 2006. Foamy virus vector integration sites in normal human cells. *Proc Natl Acad Sci U S A* 103(5):1498-1503.
- van Gent DC, Vink C, Groeneger AA, Plasterk RH. 1993. Complementation between HIV integrase proteins mutated in different domains. *The EMBO journal* 12(8):3261-3267.
- Yang F, Leon O, Greenfield NJ, Roth MJ. 1999. Functional interactions of the HHCC domain of moloney murine leukemia virus integrase revealed by nonoverlapping complementation and zinc-dependent dimerization. *J Virol* 73(3):1809-1817.
- Yang F, Roth MJ. 2001. Assembly and catalysis of concerted two-end integration events by Moloney murine leukemia virus integrase. *J Virol* 75(20):9561-9570.

- Zhao Z, McKee CJ, Kessl JJ, Santos WL, Daigle JE, Engelman A, Verdine G, Kvaratskhelia M. 2008. Subunit-specific protein footprinting reveals significant structural rearrangements and a role for N-terminal Lys-14 of HIV-1 Integrase during viral DNA binding. *The Journal of biological chemistry* 283(9):5632-5641.
- Zheng R, Jenkins TM, Craigie R. 1996. Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. *Proceedings of the National Academy of Sciences of the United States of America* 93(24):13659-13664.

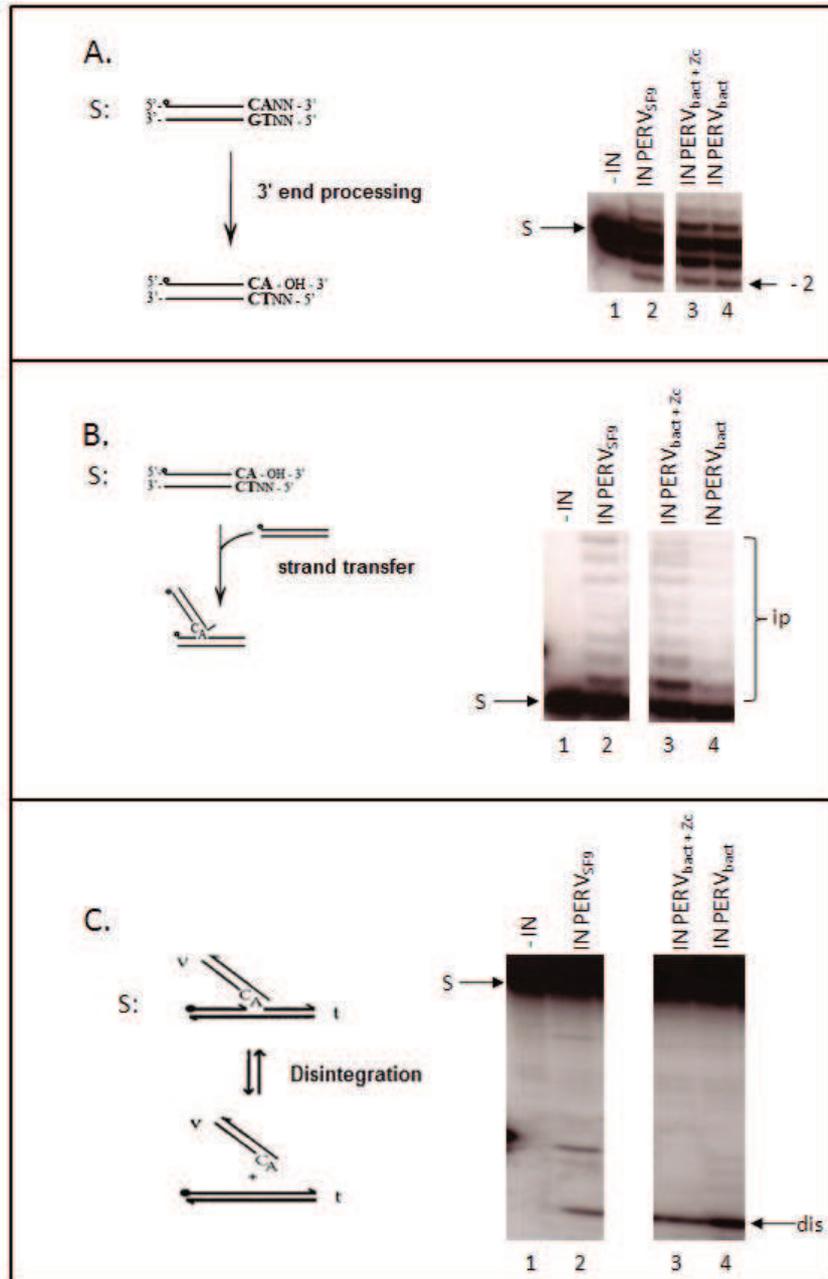
## FIGURES



**Figure 1. Purity and oligomeric state of PERV INs.**

**(A) PERV INs purified by  $\text{Ni}^{2+}$ -HiTrap ImAC HP column chromatography.** 1.5  $\mu\text{g}$  of PERV IN purified from *E. Coli* (PERV IN<sub>bact</sub>) or from SF9 cells (PERV IN<sub>SF9</sub>), were separated on a 12% SDS-polyacrylamide gel and visualized by Coomassie blue staining, before dialysis. Arrows point to a 48.2 kDa protein for His-tagged PERV IN<sub>bact</sub> and to a 47.2 kDa for His-tagged PERV IN<sub>SF9</sub>. The molecular weights of the protein standards (M) are represented on the left of panels.

**(B) Protein-protein cross-linking of PERV IN<sub>bact</sub>.** 2  $\mu\text{g}$  of PERV IN<sub>bact</sub> (-ZnCl<sub>2</sub>) and PERV IN<sub>bact zc</sub> (+ZnCl<sub>2</sub>) were incubated with or without Bis (sulfosuccinimidyl) suberate (BS<sup>3</sup>) (2 mM). Reaction products were analysed on 6% SDS-polyacrylamide gels and revealed by Western blotting using an anti-Histidine antibody. The migration of cross-linked species, monomers (M) and dimers (D) are marked. - BS<sup>3</sup>, without BS<sup>3</sup>; + BS<sup>3</sup> cross-linking of PERV IN<sub>bact zc</sub> in presence of 2 mM of BS<sup>3</sup>.



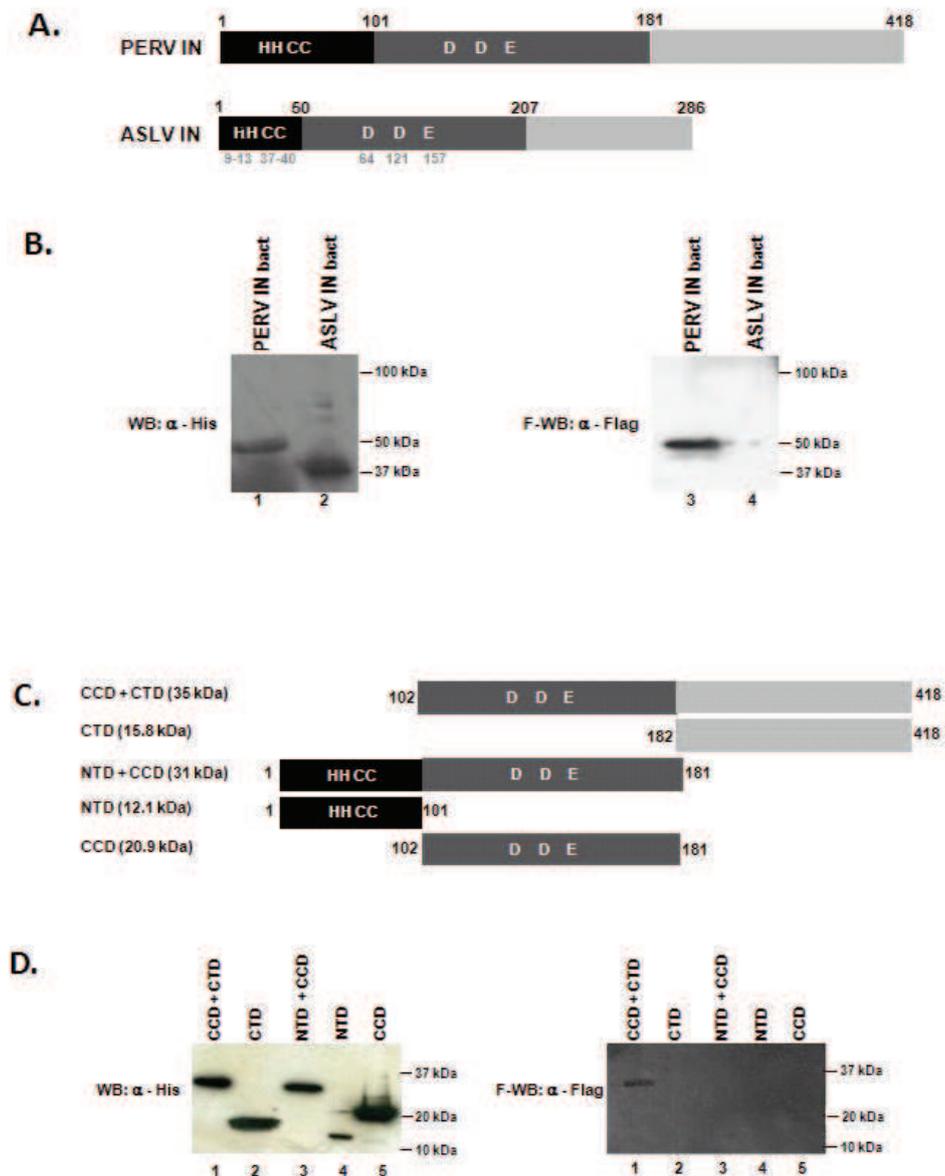
**Figure 2. Integration assays.**

**(A). 3'-processing.** Left: Illustration of the double strand oligonucleotide DNA substrate and of the 3'-processing reaction liberating a substrate processed of 2 nucleotides on the radio-labelled strand. Right: 3-processing catalyzed by PERV IN<sub>SF9</sub>, PERV IN<sub>bact\_zc</sub> and PERV IN<sub>bact</sub>. The viral DNA end oligonucleotide substrate (S) radio-labelled at the 5' end is converted into a -2 end-processed products.

**(B). Strand transfer.** Illustration of the pre-cleaved DNA substrate and of the strand transfer reaction. Strand transfer was catalyzed by PERV IN<sub>SF9</sub>, PERV IN<sub>bact\_zc</sub> and PERV IN<sub>bact</sub>. The position of the radio-labelled substrate (S) and integration products (ip) are indicated.

**(C). Disintegration.** Illustration of the Y-type DNA substrate and of the disintegration reaction. Disintegration reaction catalyzed by PERV IN<sub>SF9</sub>, PERV IN<sub>bact\_zc</sub> and PERV IN<sub>bact</sub>. The position of the radio-labelled substrate (S) and of the disintegration product (dis) are indicated.

–IN indicates the 3' processing, strand transfer and disintegration reactions performed without IN. Here are shown representative gels from two experiments.



**Figure 3. Far Western Blot**

**(A). Domain organisation of the PERV IN and ASLV IN proteins.** Representation of the His-tagged proteins used in figure 3B. The name each protein is indicated on the left.

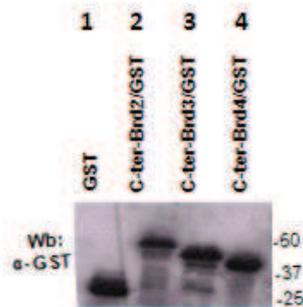
**(B). Far western blot analysis with wild type integrases.** 1 $\mu$ g of PERV IN and ASLV IN were deposited on 10 % SDS-polyacrylamide gels and transferred onto nitrocellulose membranes. The first membrane (left, wb:  $\alpha$ -His) is the result of western blot analysis using anti-His antibody, and showing the expression of the IN proteins. The second membrane was

incubated with 293T cells lysates over-expressing the 3xFlag-Brd2 protein. The membrane was washed and probed with an anti-Flag antibody (Fwb:  $\alpha$ -Flag). The positions of molecular weight markers are indicated on the right of each membrane.

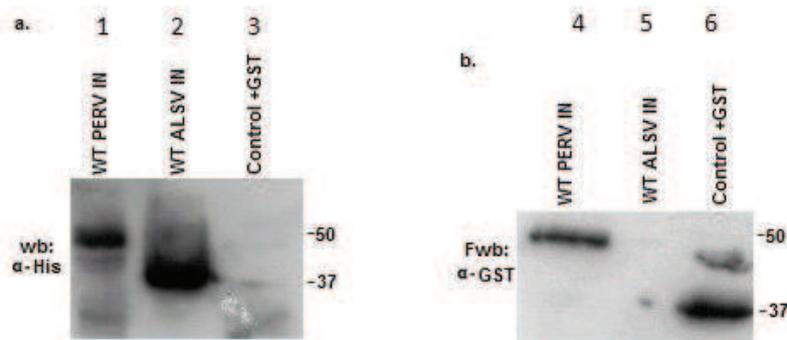
**E.**



**F.**



**G.**



**(C). Schematic representation of the one- or two-domains PERV IN used.** Delineation of the His-tagged domains of PERV IN used in figure 3D. The name and size of each mutant are indicated on the left of each protein.

**(D). Far western blot analysis with truncated IN mutants.** Cell lysates from SF9 cells expressing the His-tagged CCD + CTD (lane 1), CTD (lane 2), NTD +CCD (lane 3), NTD

(lane 4) and CCD (lane 5) domains of PERV IN were deposited on 15% SDS-polyacrylamide gels and transferred onto nitrocellulose membranes. The first membrane (left, wb:  $\alpha$ -His) is the result of western blot analysis using anti-His antibody, and showing the expression of the different domains of PERV IN from SF9 cells lysates. The second membrane was incubated with 293T cells lysates over-expressing the 3xFlag-Brd2 protein. The membrane was washed and probed with an anti-Flag antibody (Fwb:  $\alpha$ - Flag). The positions of molecular weight markers are indicated on the right of each membrane.

**(E). Domain organization of the human Brd2 and delineation of the Brd2, Brd3 and Brd4 C-terminal domain.** Schematic representation of the full length Brd2 protein, and the carboxy-terminal regions of the GST-Brd2 (aa 601 to 801), GST-Brd3 (aa 569 to 726) and GST-Brd4 (aa 607 to 722).

**(F). Expression of different C-ter-Brd proteins.** Bacterial lysates containing GST (lane 1), C-ter-Brd2/GST (lane2), C-ter-Brd3/GST (lane3), C-ter-Brd4/GST (lane4) were deposited on 15% SDS-polyacrylamide gel, then transferred onto nitrocellulose membranes. The membrane is the result of western blot analysis using anti-GST antibody and showing the expression of the C-ter domains of Brd2, Brd3 and Brd4.

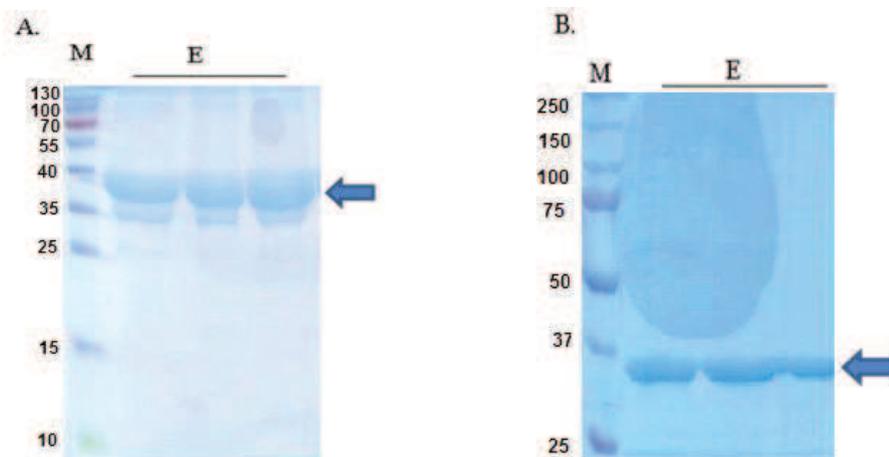
**(G). Far Westen blot analysis with the C-terminal domain of Brd2.** The first membrane (wb:  $\alpha$ -His) is the result of western blot analysis using anti-His antibody, and showing the expression of the PERV and ASLV INs. ET-Brd4/GST (GST tag control) was not detected as expected. The second membrane shows the far western blot analysis result (Fwb:  $\alpha$ - GST). In this experiment, 1 $\mu$ g of purified PERV and ALSV INs, and 15 $\mu$ l of bacterial lysate containing ET-Brd4/GST (positif control) were subjected to 10% SDS-PAGE, transferred to a nitrocellulose membrane and incubated with bacterial lysate containing the C-ter domain of Brd2. The membrane was washed and probed with anti-GST antibody. The positions of molecular weight markers are indicated on the right of each membrane.

### III- Données supplémentaires

#### A. Purification de l'IN d'ASLV

La séquence de l'IN d'ASLV a été clonée dans les plasmides pET14b et PVL1392 afin de la produire en bactérie *E. coli* souche BL21(plysS) et en cellules SF9. Les protéines comportaient une étiquette 6xhistidine en position N-terminale permettant leur purification sur une colonne de nickel immobilisé. Pour la purification des protéines IN<sub>bact</sub> et IN<sub>SF9</sub>, les surnageants des lysats ont été injectés sur une colonne HiTrap IMAC HP (GE Healthcare) chargée en nickel. La colonne a été ensuite lavée et les protéines ont été éluées par un gradient d'imidazole allant de 25 mM à 1M, collectant ainsi 0.5 ml de protéines pures par fraction. Après élution, on a observé par électrophorèse la présence d'une bande dans l'éluat migrant vers 37kDa, ce qui correspond bien à la masse attendue (figure 21). Les fractions d'INs pures ont été dialysées dans un tampon contenant 10% de glycérol, et en absence d'imidazole, puis ont été fractionnées et stockées à -80°C.

Le rendement de protéines pures obtenues par le système bactrien est plus élevé qu'à partir du système baculovirus / cellules SF9.

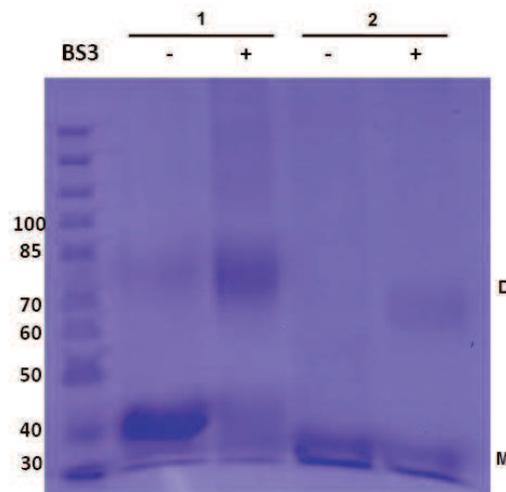


**Figure 21. Purification de l'IN d'ASLV.** Electrophorèse SDS-PAGE des fractions non retenues sur la colonne HiTrap après purification de l'IN d'ASLV produite en bactéries (figure A) ou en cellules SF9 (figure B). Les fractions éluées (E) ont été déposés sur un gel 12% (figure A) et 10% (figure B). L'IN est indiquée par une flèche. Les marqueurs moléculaires sont indiqués à gauche du gel ; les marqueurs utilisés correspondent à 5µl du « PageRuler™ Prestained Protein Ladder, Fermentas » (ref : SM0671) dans la figure A, et à 5µl du « precision plus protein standards, dual color Bio-RAD » (ref : #161-0374) dans la figure B.

### B. Cross link de l'IN d'ASLV

Comme pour la protéine IN PERV, une quantité de 2µg de protéines pures d' IN<sub>bact</sub> ASLV, ou d'IN<sub>SF9</sub> ASLV est mélangée à 2mM BS<sup>3</sup> et incubée pendant 30 minutes à température ambiante dans un tampon contenant 100mM d'Hepes, pH8 et 600mM de NaCl. La réaction est arrêtée par une incubation pendant 15 minutes à température ambiante, dans une solution « d'arrêt » contenant 25mM tris HCl, pH7.5 et 3mM glycine. Les différentes formes oligomériques sont séparées sur gel d'acrylamide 8%. Après migration à 100volts, on colore le gel au bleu de coomassie. Le résultat du cross link de l'IN d'ASLV montre deux bandes, une qui correspond à la forme monomérique de l'IN et une seconde qui correspond à un dimer d'IN (figure 22).

A l'inverse de l'IN de PERV, IN ASLV ne nécessite pas une dialyse en présence du zinc afin d'obtenir une multimérisation de la protéine.



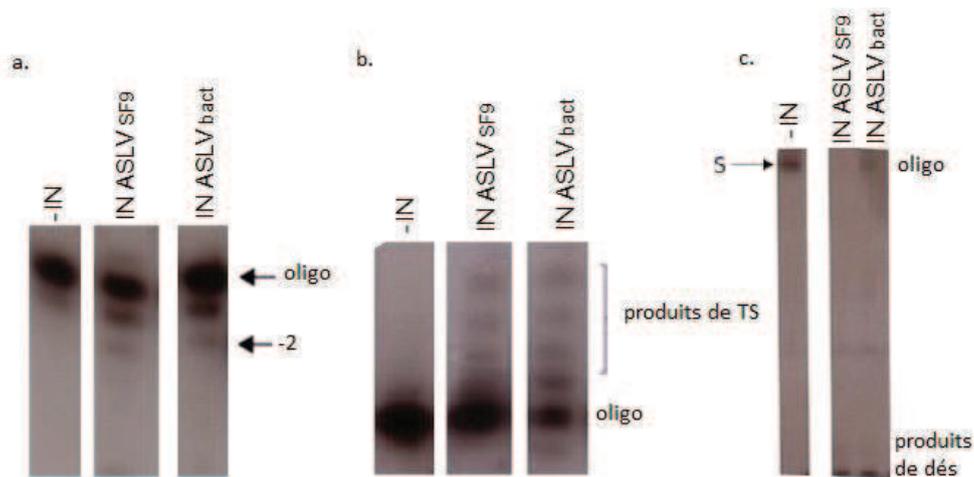
**Figure 22. Cross-link de l'IN d'ASLV.** Electrophorèse SDS-PAGE (8%) des protéines IN d'ASLV traitées ou pas avec du BS<sup>3</sup>. Les marqueurs moléculaires sont indiqués à gauche du gel. 1: IN<sub>bact</sub> ASLV, 2: IN<sub>SF9</sub> ASLV, M: monomère, D: dimère.

### C. Activités catalytiques de l'IN d'ASLV

Les bactéries sont couramment utilisées dans les protocoles de production de l'IN d'ASLV au sein de notre laboratoire à Lyon. Dans une tentative d'améliorer les tests d'activité de la protéine, l'IN d'ASLV a été donc produite dans des cellules SF9, capables de lui procurer des modifications post-traductionnelles. Les réactions de clivage, de transfert de

brins et de désintégration, réalisées par les deux enzymes  $IN_{bact}$  ASLV et  $IN_{SF9}$  ASLV ont été testées *in vitro*, selon les conditions décrites dans la partie ‘matériels et méthodes’. Or, dans la fonctionnalité de cette enzyme en comparaison avec celle produite en système bactérien.

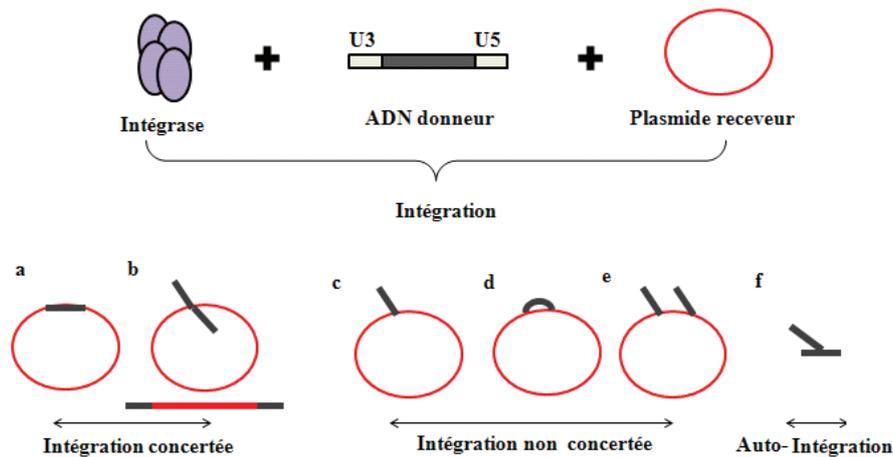
Contrairement à nos attentes, les résultats des tests *in vitro* de l’activité de la protéine produite en système baculovirus / cellules SF9 ne montrent aucune amélioration en comparaison avec l’intégrase produite en système bactérien (figure 23).



**Figure 23. Activités catalytiques de l’IN d’ASLV *in vitro*.** (A) clivage : réaction de clivage réalisée par l’IN bact, ou l’IN baculo, produisant un produit clivé de -2 nucléotides. (B) transfert de brins : représentation du résultat du transfert de brins réalisé par les IN bact et baculo. Sont également indiqués les oligos et les produits du transfert de brins. (C) désintégration : résultats du test de désintégration réalisé par l’IN bact et IN baculo montrant les oligonucléotides marqués et les produits de désintégration. -IN indique les réactions de clivage, transfert de brins et désintégration en absence de l’IN.

#### D. Test de l’intégration concertée de PERV

Les données obtenues ici à partir des tests d’activités catalytiques *in vitro* de PERV, correspondent au mécanisme d’intégration d’une seule des extrémités virales. Or, *in vivo*, l’intégration correspond à l’insertion simultanée des deux extrémités virales sur un même site de l’ADN cellulaire, on parle alors d’intégration concertée. La réaction d’intégration concertée peut être reproduite *in vitro*; la reconstitution des conditions optimales permet de comprendre ce qui se passe *in vivo*.



**Figure 24. Représentation de la réaction d'intégration *in vitro* et des différents produits obtenus.** Les produits d'intégration concertée sont désignés par a et b. Ils correspondent à l'insertion simultanée des deux extrémités d'un ADN donneur, ou de deux ADNs donneurs différents, sur un même site de l'ADN receveur. Les produits d'intégration non concertée sont désignés par c, d, e. Ils correspondent à l'insertion d'une seule extrémité d'un ADN donneur dans l'ADN receveur, ou l'insertion de deux extrémités d'un, ou de deux ADNs donneurs, sur un deux sites différents de l'ADN receveur. Les produits d'auto-intégration sont désignés par f, et correspondent à l'insertion d'un ADN donneur dans un autre.

Afin de reproduire la réaction d'intégration réalisée par l'intégrase de PERV telle qu'elle est observée *in vivo*, la protéine intégrase de PERV active et pure est incubée avec un ADN donneur représentant l'ADN viral et un plasmide receveur représentant l'ADN de la cellule cible.

L'ADN donneur utilisé durant ce test correspond à un ADN pré-clivé, mimant les extrémités virales après clivage par l'intégrase, et comportant le gène *supF* (250 pb) et 20 pb des séquences U3 et U5 du virus PERV. L'obtention de l'ADN pré-clivé se fait par digestion du plasmide pUC19 *supF* PERV par l'enzyme NdeI. Le plasmide receveur utilisé est pBSK- $\Delta$ 2-zeo. Les conditions de la réaction ont alors été adaptées à partir d'un test d'intégration concertée du modèle MLV (Yang F and Roth MJ 2001). Ainsi, 2 $\mu$ g d'IN de PERV sont incubées une nuit à 4°C avec 1 $\mu$ g de plasmide receveur et 100ng de plasmide donneur. Le lendemain, 100 $\mu$ l de tampon d'intégration (20 mM MES PH6.2, 10 mM DTT, 10 mM MnCl<sub>2</sub>, 10% DMSO, 200 mM KCl, 10% glycerol) est ajouté, et le mélange est incubé pendant 1h30' à 37°C. Les produits d'intégration sont ensuite digérés avec l'enzyme BamHI qui présente un site unique de restriction au niveau de l'ADN donneur *supF*, puis incubés avec la T4 DNA ligase. Ainsi, on récupère toutes les formes d'intégration concertée, les produits de forme 'a', et ceux de formes 'b' re-circularisés (figure 24).

Ensuite, les produits d'intégration vont servir à transformer 100 µl de bactéries MC1061/P3 par choc thermique. Les bactéries sont enfin sélectionnées sur boîte LB contenant un mixe de 4 antibiotiques (ampicilline, zéocine, tétracycline et kanamycine). Le gène de résistance à la zéocine se trouve dans le plasmide pBSK- Δ2-Zéo. Les gènes de résistance pour l'ampicilline, la kanamycine et la tétracycline se trouvent dans l'épisode P3 des bactéries MC1061/P3. Les gènes ampicilline et tétracycline sont porteurs d'une mutation *amb*. Ainsi, seules les bactéries ayant intégré un plasmide porteur du gène *supF* sont capables de se multiplier en présence des 4 sélections antibiotiques. Ces bactéries contiennent donc les produits d'intégration dont les deux extrémités d'un ADN donneur ont été insérées dans un plasmide receveur.

Les tests réalisés jusqu'à présent n'ont permis d'obtenir qu'une seule colonie bactérienne sur la boîte gélose en présence des 4 sélections antibiotiques. Il est donc nécessaire d'optimiser les conditions du test. Plusieurs modifications peuvent être rapportées afin d'améliorer les résultats: (i) tester une gamme de différentes concentrations d'IN. (ii) tester différents tampons d'intégration. (iii) modifier le temps de la réaction d'intégration. (iv) tester un ADN donneur comportant des bouts francs mimant les extrémités virales avant action de l'IN.

#### *E. Interaction INs-TAB182*

Dans l'article, on présente les résultats de l'interaction entre les INs et Brd2. Or, notre projet consistait à vérifier également l'interaction avec une seconde protéine cellulaire identifiée durant le double hybride mené pour l'IN de JSRV, la protéine TAB182 (182 kDa Tankyrase-1-Binding protein).

TAB182 est le produit du gène TNKS1BP1. Elle est aussi nommée KIAA1741. La fonction de TAB182 reste largement inconnue. Des données obtenues par des analyses d'immunofluorescence montrent que cette protéine peut être localisée dans le noyau et dans le cytoplasme (Seimiya H and Smith S 2002). TAB182 cytoplasmique co-localise avec l'actine corticale. Dans le noyau, TAB182 reste associée à la chromatine tout au long de la mitose. TAB182 a été également identifiée par double hybride comme un des 3 partenaires du domaine ANK de la Tankyrase 1 et 2 (Seimiya H and Smith S 2002). Sachant que la Tankyrase 1 est un membre de la famille de poly(ADP-ribose) polymérase (PARPs). Son domaine ANK est un motif répété de 33 acides aminés impliqué dans l'interaction protéine-

protéine, qui sert apparemment à recruter des substrats pour le domaine PARP. Ainsi, TAB182, IRAP (Insulin-Responsive Amino Peptidase) et TRF1 (Telomeric DNA-binding protein) possèdent en commun un motif RXXPDG qui joue un rôle essentiel dans leur interaction avec les Tankyrase1 et 2 (Sbodio JI and Chi NW 2002).

Pour vérifier s'il existe une interaction entre TAB182 et les INs de PERV et d'ASLV, on a utilisé la méthode 'Far western blot'. Ainsi, 2 µg de protéines IN PERV et IN d'ASLV pures ont été transférées sur membrane nitrocellulose, qui a été ensuite incubée avec 300µl de lysat de cellules 293T exprimant la protéine TAB182 fusionnée à une étiquette 3xFlag. Après une nuit d'incubation, la membrane a été lavée et les protéines sont révélées avec de l'anticorps anti-Flag. Si une interaction existait entre une IN et TAB182, on devrait obtenir une bande correspondante à la taille de l'IN. Or, le résultat de l'expérience montre qu'il n'existe aucune interaction entre l'IN de PERV et TAB182, ni entre l'IN d'ASLV et TAB182 (résultats non montrés). Cependant, en absence d'un contrôle positif (une protéine interagissant avec TAB182), ces résultats restent à confirmer.

#### **IV- Discussion des résultats**

L'étape principale du cycle viral, l'intégration, est réalisée par une enzyme virale, l'intégrase. Plusieurs protéines cellulaires participent au processus d'intégration en interagissant avec l'intégrase. Le blocage de chacune de ces interactions constitue une cible potentielle pour le développement d'antirétroviraux efficaces. Actuellement, la cible la plus prometteuse concerne l'interaction entre l'IN de HIV-1 et LEDGF/p75. Dans le présent projet, nous nous sommes intéressés à deux modèles de rétrovirus (PERV et ASLV) dans le but principal d'identifier des partenaires cellulaires de leurs intégrases. Une telle étude nécessitait la production de ces deux protéines en état d'activité. La production de l'intégrase d'ASLV pure et active se faisait par des méthodes classiquement utilisées au laboratoire. Par contre, aucune donnée n'était jusqu'à présent disponible sur l'intégrase de PERV, ce qui nous a poussés à approfondir la caractérisation de cette protéine en testant son activité *in vitro*. Différents paramètres de purification et de tests d'activités *in vitro* ont été donc mis au point, ainsi qu'une technique de validation d'interactions protéine-protéine, appelée far western blot qui a permis de confirmer et de caractériser l'interaction entre l'IN de PERV et la protéine cellulaire Brd2.

##### *A. Comparaison des rendements de protéines pures obtenues par les systèmes bactérien/baculoviral*

Le premier objectif de cette étude était l'acquisition de nouvelles données fonctionnelles et structurales sur l'IN de PERV. Pour répondre à cet objectif, il fallait au préalable développer une méthode de purification de l'IN permettant d'obtenir à la fois un rendement élevé mais également une protéine fonctionnelle. Afin de produire des protéines recombinantes identiques aux protéines naturelles, nous avons décidé d'utiliser, en comparaison avec le système bactérien habituel, des cellules d'insectes comme système de production.

A l'issue des expériences de purification réalisées à partir d'une production de protéines en système bactérien ou baculovirus / cellule d'insecte, le protocole de purification adapté de Roth et al, nous a permis d'obtenir une protéine en quantité suffisante, et en concentration satisfaisante pour la réalisation des tests d'activité *in vitro*. En effet, les concentrations obtenues pour les INs de PERV et d'ASLV sont de 4 mg et 2 mg respectivement, pour les

protéines produites dans 1 litre de culture bactérienne, et de 2 mg et 1.5 mg pour les protéines produites dans 1 litre de culture de cellules d'insecte.

Bien que les avantages du système d'expression en baculovirus / cellules d'insecte soient nombreux, notamment sa capacité d'incorporer des modifications post-traductionnelles souvent essentielles à l'activité biologique des protéines synthétisées, des problèmes ont été rencontrés lors de l'utilisation de ce système. En effet, il y a eu une précipitation des protéines durant l'extraction des cellules d'insectes. La solubilisation par un détergent a été donc nécessaire afin d'extraire aisément la protéine d'intérêt et de la maintenir en solution. Quant aux protéines produites en système bactérien, l'extraction des protéines a été réalisée en absence de détergent.

### *B. Activités catalytiques des intégrases*

Les activités catalytiques des intégrases ont été évaluées par un ensemble de tests réalisés *in vitro* : cross-link, clivage, transfert de brins, désintégration. Nos résultats montrent que (i) l'IN produite en système bactérien présente une activité plus élevée que celle produite en cellules d'insectes, (ii) le cation divalent zinc joue un rôle important dans l'activité de l'IN de PERV, notamment dans l'oligomérisation de la protéine qui en est dépendante.

#### *a) Effet du détergent*

Plusieurs facteurs peuvent affecter l'activité de la protéine, parmi lesquels sa pureté et sa conformation. Or, les intégrases produites en bactérie et cellules SF9 semblent présenter un niveau de pureté équivalent. Cependant, l'addition du détergent dans le tampon d'extraction de protéines des cellules SF9 aurait pu causer une dénaturation de la protéine, se traduisant pas par une baisse de son activité.

Bien que les détergents soient couramment utilisés pour libérer les protéines recombinantes en dissolvant les membranes du compartiment cellulaire dans lequel elle se trouve, leur emploi cependant doit être fait de façon contrôlée car ils peuvent briser les lysosomes, ce qui libère des enzymes hydrolytiques (protéases, nucléases, etc.) qui peuvent attaquer et détruire les protéines. Pour éviter un tel problème, nous avons inclus dans le tampon de lyse un mélange d'inhibiteurs de protéases à large spectre d'action. De plus, nous avons choisi un détergent non ionique, particulièrement peu dénaturant, afin de préserver

l'activité de la protéine. Malgré toutes les précautions prises, une dénaturation causée par l'ajout de détergent reste la seule explication au fait que la protéine présente un niveau d'activité faible.

#### *b) Rôle du zinc*

La capacité de l'intégrase de PERV à dimériser, semblent être totalement dépendante de la présence des ions  $Zn^{2+}$ ; en effet, l'absence de  $ZnCl_2$  dans le tampon de dialyse de cette protéine enlève toute possibilité de formation de dimère durant le test de « cross-link ». Ceci pourrait être expliqué par le fait que le domaine N-terminal de l'intégrase comporte un motif HHCC analogue à un doigt à zinc fixant effectivement ce cation et dont le rôle pourrait être de favoriser le processus de multimérisation de la protéine entière. En effet, Zheng et ses collègues ont démontré que la fixation du zinc par le motif HHCC induit le repliement du domaine N-terminal de l'intégrase de HIV-1. En solution, ce domaine existe sous forme d'un équilibre monomère-dimère, ce qu'a confirmé la résolution de sa structure par résonance magnétique nucléaire. Chaque unité monomérique est structurée en quatre hélices  $\alpha$ , les hélices  $\alpha_2$  et  $\alpha_3$  formant un motif hélice-tour-hélice, stabilisées par la coordination du zinc et par des interactions hydrophobes (Zheng R, Jenkins TM et al. 1996). De même, Roth et ses collègues ont démontré qu'en présence du zinc, les domaines HHCC de l'intégrase de MLV forment un dimère stable en solution (Yang F, Leon O et al. 1999). Plus récemment, Pandey et ses collègues ont montré par gel filtration que l'IN du VIH-1 pure est présente sous forme d'un tétramère en présence de  $50\mu M$  de  $ZnCl_2$ , alors qu'elle se trouve sous forme monomérique en absence du zinc (Pandey KK, Bera S et al. 2011). Les résultats obtenus par notre laboratoire sont donc en corrélation avec les données obtenues sur le rôle du zinc dans la dimérisation des intégrases d'autres rétrovirus.

Un effet stimulateur du zinc sur l'activité de l'intégrase de PERV se fait également ressentir dans les tests d'activités catalytiques *in vitro*. Les deux réactions, 3'-processing et transfert de brins, réalisées par l'intégrase de PERV sont plus efficaces quand la protéine est renaturée en présence du zinc. Notons que la présence du cation  $Zn^{2+}$  n'est pas requise pour la multimérisation ni l'activité de l'IN d'ASLV *in vitro*.

### *C. Spécificité de l'interaction IN PERV/ Brd2*

Le deuxième objectif de cette étude était la validation d'interactions potentielles entre les protéines d'intérêt (IN d'ASLV, IN de PERV) et la protéine cellulaire (Brd2). Nous avons pu examiner *in vitro* ces interactions par une méthode appelée far western blot.

Le far western se différencie des techniques plus classiques qui sont le double hybride, le pull down, la co-immunoprécipitation; l'avantage d'une telle technique est qu'elle permet d'identifier des partenaires directs, contrairement aux techniques précédemment citées qui permettent la détection de partenaires indirects au sein de complexes macroprotéiques. Or, cette approche présente également des limites. Par exemple, les protéines peuvent être dénaturées et dans ce cas l'interaction ne sera pas détectée par la méthode far western. En effet, les protéines sont migrées sur des gels de polyacrylamide en présence de détergents ioniques (SDS) et d'agents réducteurs ( $\beta$ -mercaptoéthanol, dithiothréitol) qui dénaturent les protéines. Les agents dénaturants sont généralement déplacés lors de l'électro-transfert ce qui restitue à beaucoup de protéines une activité plus ou moins partielle. Aussi, un facteur déterminant pour détecter avec succès une interaction protéine-protéine sur un support solide est la renaturation de la protéine sur son support.

Dans notre étude, on cherchait à dépister la capacité de Brd2 présente dans un extrait cellulaire total à interagir avec les protéines IN d'ASLV et de PERV pures transférées sur une membrane nitrocellulose. Cette méthode nous a permis de mettre en évidence une interaction directe entre l'IN de PERV et Brd2. En revanche, aucune interaction n'a été observée par cette méthode entre l'IN d'ASLV et Brd2. Le résultat n'enlève pas la possibilité d'existence d'une interaction indirecte entre l'IN d'ASLV et Brd2, ni même permet d'assurer que l'absence d'interaction n'est pas due à une dénaturation de la protéine après son passage sur un gel SDS-PAGE. Ces données restent donc à confirmer par une autre technique, plus classique, de validation d'interaction protéine-protéine.

### *D. Caractérisation de l'interaction entre les domaines de l'IN PERV et Brd2*

Dans la mesure où nous avons à notre disposition les outils nécessaires qui permettent d'avoir une vision plus précise quant à l'interaction de l'IN de PERV avec Brd2, nous avons testé par la même méthode l'interaction de Brd2 avec les différents domaines de l'IN PERV. Nos résultats montrent qu'aucun des domaines séparés ne peut interagir avec l'intégrase, et que seule la combinaison des deux domaines catalytique central/C-terminal pouvait interagir

avec Brd2. Nous avons également démontré par far western blot l'implication du domaine C-terminal de Brd2, préalablement défini comme un domaine d'interaction protéine-protéine. De plus, nos résultats étaient en corrélation avec ceux obtenus par Roth et ses collègues pour l'IN de MLV ; en effet, durant le double hybride réalisé pour l'intégrase de MLV contre une banque d'ADNc murins, 9 séquences de Brd2 interagissant avec l'IN de MLV ont été isolées, dont 8 correspondent au domaine C-terminal de Brd2. Les données obtenues pour l'IN de MLV montrent également que, seule la combinaison des deux domaines catalytique central et C-terminal peut interagir avec Brd2 (Studamire B and Goff SP 2008).

Par ailleurs, nous avons voulu savoir si cette interaction s'étendait à d'autres protéines de la même famille de Brd. Pour cela, un far western blot a été réalisé avec l'intégrase comme cible à reconnaître par le domaine C-terminal de Brd3 ou Brd4 présent dans un extrait bactérien. Ainsi, aucune interaction n'a été démontrée avec ces domaines.

## **CONCLUSION GENERALE ET PERSPECTIVES**

## **I- Conclusion générale**

Les études fonctionnelles de l'intégrase sont importantes pour acquérir une meilleure compréhension du fonctionnement de cette enzyme. Nos résultats ont permis d'apporter un nouveau modèle d'étude des intégrases des gammarétrovirus, qui est l'intégrase de PERV, et de confirmer l'importance du cation zinc dans la multimérisation de cette enzyme.

De plus, l'étude systématique des sites d'intégration montre que l'intégration des rétrovirus n'est pas totalement aléatoire. Dans le cas du PERV, les études des sites d'intégration *in vivo* ont montré que les sites d'initiation de la transcription, ainsi que les îlots CpG sont favorisés pour l'intégration. Bien que les structures locales de l'ADN cible (courbure, flexibilité) puissent jouer un rôle, il est vraisemblable que ce biais d'intégration est également dû à l'interaction des CPI avec des composants d'origine cellulaire. Le travail présenté ici a permis d'identifier le premier facteur cellulaire capable d'interagir avec l'IN de PERV. Le fait que Brd2 soit un facteur de transcription suggère que celui-ci pourrait être effectivement responsable du ciblage de l'intégration. Étant donné que Brd2 interagit directement avec l'intégrase, le domaine d'interaction de ces deux protéines constituera une nouvelle cible séduisante pour des ligands de l'intégrase à activité antivirale.

Enfin, l'identification d'un tel cofacteur de l'intégrase permet d'avoir une vue plus précise des déterminants de l'intégration *in vivo* et ouvre de nouvelles perspectives de recherche pharmacologique.

## **II- Perspectives expérimentales**

### *A. Optimisation des protocoles de production et de purification de l'intégrase*

Deux systèmes de production de protéines ont été utilisés dans ce travail pour exprimer l'IN de PERV et d'ASLV. Or, ces deux systèmes présentent des inconvénients. Un des principaux inconvénients d'*E. coli* est qu'elle sécrète mal les protéines, et qu'elle soit incapable de réaliser les modifications post-traductionnelles de type eucaryote. Ceci peut avoir des conséquences essentielles sur la conformation ou l'activité biologique de la protéine. Le second système utilisé dans ce travail est le système baculovirus/cellules d'insecte, capable de réaliser la plupart des modifications post-traductionnelles courantes (formation de ponts disulfures, oligomérisation, acétylation, phosphorylation, l'insertion membranaire, etc.). Cependant, ce système a l'inconvénient d'être lytique et donc transitoire. Le caractère lytique

du système et les très hauts niveaux d'expression ont pour conséquence que les capacités de maturation et de sécrétion des cellules sont engorgées, ce qui explique en partie que les modifications post traductionnelles se font parfois de façon moins efficace.

Afin d'optimiser l'activité des intégrases, il serait donc envisageable d'exprimer ces protéines dans d'autres systèmes de production. Les cellules mammifères, par exemple, permettent de générer des protéines avec des modifications post traductionnelles parfaites, mais leur rendement de production est plus faible que les autres systèmes et leur vitesse de croissance est plus longue, ce qui explique le fait qu'elles ne soient pas utilisées pour la production des intégrases à haut débit. Par contre, l'utilisation de la levure est envisageable puisqu'elle permet de produire des protéines recombinantes avec des modifications post-traductionnelles proches de celles des eucaryotes supérieurs, le rendement est meilleur, le délai d'expression est plus court, de plus que l'efficacité de ce système a été prouvée par Vera J et ses collègues pour la production de l'IN de MLV (Vera 2005). En effet, cette protéine présentait une activité deux fois plus élevée que celle purifiée à partir d'une production en bactérie.

Notons que d'autres modifications peuvent être rapportées, directement au mode de purification, de sorte à augmenter la pureté et l'activité de la protéine. En effet, l'utilisation des protéines sans étiquette histidine, ou le clivage de cette dernière après purification pourrait, peut être, améliorer l'activité de la protéine. D'autre part, une purification sélective des formes tétramériques de l'intégrase par gel filtration pourrait permettre la récupération d'une forme spécifique de l'IN à priori active pour l'intégration (Pandey KK, Bera S et al. 2011).

### *B. Validation de l'interaction IN PERV-Brd2*

Si la mise en place d'un test *in vitro* nous a apporté quelques informations sur l'interaction IN PERV/Brd2, de nombreux travaux restent à optimiser afin de confirmer nos résultats.

Dans un premier temps, il serait intéressant de valider l'interaction entre l'IN PERV et Brd2 par des techniques plus courantes, comme le double hybride ou co-immunoprécipitation, et par co-localisation en microscopie confocale des deux protéines partenaires dans les compartiments cellulaires. La technique de Far western blot étant *in vitro*, le double hybride permettra donc de valider l'interaction IN PERV/Brd2 en levure. Le protocole de double

hybride envisagé sera basé sur la méthode décrite par Studamire et ses collègues pour la recherche des cofacteurs de l'IN de MLV (Studamire B and Goff SP 2008).

Or, la technique double hybride présente certains inconvénients, comme la présence de nombreux faux-positifs et faux-négatifs. En effet, il est possible par exemple d'induire de fausses interactions causées par la surexpression des protéines. Quant à la présence de faux négatifs, elle peut être causée, entre autres, par une structure tertiaire erronée des protéines, provoquée par des modifications post-traductionnelles modifiées ou absentes, ou un repliement incorrect. Il serait donc intéressant de confirmer cette interaction par une autre méthode, comme la co-localisation des deux protéines IN PERV et Brd2, marquées par immunofluorescence, dans des cellules humaines par microscopie confocale. Des premiers essais ont été réalisés en utilisant le microscope confocal spectral Leica TCS-SP2, disponible sur la plateforme SFR BioSciences Gerland - Lyon Sud. Les résultats obtenus jusqu'à présent ne permettent pas de confirmer la co-localisation de l'intégrase avec le cofacteur Brd2. Le test de différentes conditions est actuellement en cours afin d'optimiser la production des deux protéines, et améliorer le marquage pour faciliter la détection des protéines par microscopie.

### *C. Evaluation du rôle de Brd2 dans l'intégration*

Le fait que Brd2 soit capable de se fixer d'une part à la chromatine acétylée grâce à ses bromodomaines, et d'autre part à l'intégrase grâce à son domaine C-terminal, nous permet de suggérer qu'elle pourrait avoir un rôle semblable à celui de LEDGF/p75 dans l'intégration rétrovirale.

Pour cela, on envisage dans un premier temps de tester les activités catalytiques de l'IN de PERV *in vitro* (clivage, transfert de brins, désintégration) comme décrit précédemment, mais en présence du domaine C-terminal de Brd2, afin de tester si Brd2 serait capable de stimuler le mécanisme d'intégration. Notons que le substrat utilisé généralement au cours de ces tests correspond à un ADN nu et ne représente pas la situation *in vivo* où l'ADN cible se trouve sous forme de chromatine. Il serait donc envisageable d'optimiser nos tests d'intégration en utilisant un substrat de polynucléosome reconstitué.

Dans un second temps, on envisage d'identifier par mutagenèse dirigée les résidus de l'intégrase impliqués dans la fixation à Brd2, ainsi que les résidus du domaine C-terminal de Brd2 impliqués dans la liaison à l'intégrase. Ainsi on pourra déterminer s'il existe une

corrélation entre la perte d'interaction avec Brd2 et une diminution de la capacité du virus à se répliquer.

En complément de ces expériences de mutagenèse, d'autres approches permettront d'établir si Brd2 est nécessaire à la réplication virale. L'approche ARN interférence ainsi que l'extinction intensive de l'expression de Brd2 permettront de souligner l'importance de son rôle sur la réplication virale.

De plus, la surexpression d'une forme dominante négative de Brd2 permettra de sélectionner des virus résistants à la mutation de l'IN. L'affinité réduite de l'intégrase mutée pour Brd2, confirmera si le virus résistant est dépendant de Brd2 pour répliquer.

Par ailleurs, le fait que Brd2 soit capable de se fixer à la chromatine acétylée via son bromodomaine, nous permet de suggérer que Brd2, comme LEDGF/p75, serait capable de cibler l'intégrase sur les chromosomes de la cellule hôte. On envisage d'abord de comparer par microscopie confocale la localisation de l'IN dans des cellules humaines ou porcines, en présence (surexpression) ou absence de Brd2 (ARN interférence).

D'autre part, l'analyse des sites d'intégration de PERV dans des cellules humaines et porcines partiellement déplétées en Brd2, ou dans des cellules KO pour le gène codant pour la protéine Brd2, confirmera si Brd2 est capable de diriger l'intégration du virus dans des sites spécifiques du génome. Les préférences d'intégration envers les gènes, les îlots CpG ou les sites d'initiation de la transcription seront comparées aux données obtenues précédemment pour l'intégration de PERV, sachant que la méthode d'analyse est maîtrisée dans notre laboratoire à l'ANSES (Moalic Y, Blanchard Y et al. 2006).

Il serait également intéressant de comparer l'effet de Brd2 et de LEDGF/p75 sur la spécificité de l'intégration rétrovirale. En effet, LEDGF/p75 est un cofacteur spécifique de l'IN des lentivirus, et le responsable principal du profil d'intégration du VIH-1, alors que Brd2 a été identifié, jusqu'à présent, comme un facteur commun aux gammarétrovirus (PERV et MLV) et lentivirus (VIH-1) ; il serait donc envisageable d'étudier la sélection des sites d'intégration de PERV en présence d'une protéine chimère, constituée du domaine C-terminal de Brd2 couplé au domaine de la liaison à la chromatine de LEDGF/p75 (PWWP et crochets AT). On vérifie ainsi si le domaine responsable de la liaison à la chromatine de LEDGF/p75 pourrait conférer à PERV un profil d'intégration spécifique des lentivirus. Sachant qu'une telle étude menée par Silvers R et ses collègues a été réalisée pour VIH-1, et a permis de

montrer que la surexpression d'une protéine chimère constituée de l'IBD de LEDGF/p75 et du chromodomaine de la protéine HP1a (heterochromatin protein-1a) et nommée TIHPLE, réduit l'intégration du VIH-1 dans les gènes (Silvers RM, Smith JA et al. 2010).

Enfin, notons que jusqu'à présent, aucune étude structurale n'a été réalisée pour l'intégrase de PERV. A part son rôle potentiel dans la réplication virale, l'identification de Brd2 comme partenaire de l'IN de PERV, pourrait faciliter l'obtention de données structurales (conformation, stabilité, oligomérisation, résidus critiques pour l'interaction) sur cette enzyme virale, seule, ou au sein de l'intasome, en s'appuyant sur les expériences menées récemment sur l'IN du VIH-1 en présence de LEDGF/p75.

### **III- Perspectives à long terme**

#### *A. Thérapies antirétrovirales*

Ce projet s'inscrit dans une démarche visant à identifier de nouveaux partenaires cellulaires de l'intégrase, essentiels à l'intégration, et de développer des molécules dirigées contre ces derniers. Actuellement, les inhibiteurs des interactions protéines virales-protéines cellulaires sont considérés comme une option très prometteuse pour contourner les mutations au sein des enzymes virales dans le cadre d'une thérapie antirétrovirale (Debyser Z and Christ F 2010; Hazuda DJ 2010). Dans notre travail, nous démontrons que Brd2 interagit avec l'IN de PERV mais pas avec l'IN d'ASLV. Notre résultat permet de suggérer que cette protéine n'est pas un partenaire commun à tous les rétrovirus, et peut donc jouer un rôle spécifique dans l'intégration de certains d'entre eux. De plus, s'il s'avère que ce régulateur de la transcription joue un rôle dans la sélection de sites spécifiques du génome de l'hôte, la conception d'inhibiteurs visant cette interaction serait très intéressante.

Pour ce faire, la zone précise d'interaction entre les domaines catalytique centrale/ C-terminal de l'intégrase de PERV et le domaine C-terminal de Brd2 devra être caractérisée. Des travaux de mutagenèse dirigée seront donc utiles afin d'identifier les résidus critiques pour l'interaction entre ces deux protéines. Une étude structurale par cristallographie des domaines d'interaction pourra être également envisagée. Ensuite, de nouveaux peptides mimant cette zone d'interaction seront créés par modélisation moléculaire. Leur activité inhibitrice sera évaluée par un test d'intégration *in vitro*. Une autre approche qu'on pourra prendre en considération est la technique de « phage display ». Celle-ci permettra de sélectionner à partir d'une banque de phages présentant des peptides, ceux qui ont une forte affinité pour l'intégrase. Une fois identifiés, ces peptides seront analysés *in vitro* dans les tests d'activités. Ceux qui présenteront un pouvoir inhibiteur important seront testés en cellules afin de déterminer s'ils disposent d'un pouvoir inhibiteur sur la réplication du virus.

#### *B. Prévention contre l'intégration du PERV et xénotransplantation*

L'utilisation du porc en xénotransplantation a soulevé des préoccupations sur les risques du transfert du gène codant PERV. Bien qu'on ait retrouvé de l'ADN de PERV chez

des receveurs humains, cependant aucun signe de virémie ou d'infection n'a été identifié. Or, les progrès récents destinés à surmonter le rejet immunitaire pourront augmenter le risque de transmission de PERV. Ce problème est actuellement surmonté par des modifications des produits de xénotransplantation de façon à diminuer la libération des virions de PERV par EOPS (élevage d'animaux exempts d'organismes pathogènes spécifiques).

Durant les dernières années, une nouvelle expertise concernant l'intégration du PERV a été acquise au sein de l'unité de recherche à l'ANSES. Précédemment, le laboratoire s'est intéressé au profil d'intégration de ce rétrovirus, qui semble être similaire à celui du MLV. Ce résultat permet de suggérer qu'un risque d'activation oncogénique par mutagenèse insertionnelle, équivalent à celui du MLV, existe en cas de transmission du PERV à l'homme. Il serait donc important d'évaluer les risques que pourrait provoquer le PERV en cas de franchissement accidentel de la barrière immunologique. De ce fait une étude plus approfondie de l'intégration de PERV était à notre avis nécessaire. Dans ce travail, nous avons recueilli des données sur l'IN de PERV, notamment sur les conditions requises pour son activité. Notre projet de recherche de partenaires cellulaires de l'intégrase contribuera également à une meilleure compréhension des conditions d'intégration de PERV, surtout si les résultats de nos études en cours révèlent, comme attendu, un rôle de Brd2 dans la sélection des sites d'intégration de ce rétrovirus.

## REFERENCES

## A

- Acel A, Udashkin BE, et al. (1998). "Efficient gap repair catalyzed in vitro by an intrinsic DNA polymerase activity of human immunodeficiency virus type 1 integrase." *J Virol* **72**(3): 2062-71.
- Agapkina J, Smolov M, et al. (2006). "Probing of HIV-1 integrase/DNA interactions using novel analogs of viral DNA." *J Biol Chem* **281**(17): 11530-40.
- Akiyoshi DE, Denaro M, et al. (1998). "Identification of a full-length cDNA for an endogenous retrovirus of miniature swine." *J Virol* **72**(5): 4503-7.
- Albanese A, Arosio D, et al. (2008). "HIV-1 pre-integration complexes selectively target decondensed chromatin in the nuclear periphery." *PLoS One* **3**(6).
- Alexander BL, Ali RR, et al. (2007). "Progress and prospects: gene therapy clinical trials (part 1)." *Gene Ther* **14**(20): 1439-47.
- Allouch A, Di Primio C, et al. (2011). "The TRIM family protein KAP1 inhibits HIV-1 integration." *Cell Host Microbe* **9**(6): 484-95.
- Andrake MD, Sauter MM, et al. (2008). "Nuclear import of Avian Sarcoma Virus integrase is facilitated by host cell factors." *Retrovirology* **5**(73).
- Andrake MD and Skalka AM (1996). "Retroviral integrase, putting the pieces together." *J Biol Chem* **271**(33): 19633-6.
- Antony JM, Deslauriers AM, et al. (2011). "Human endogenous retroviruses and multiple sclerosis: innocent bystanders or disease determinants?" *Biochim Biophys Acta* **1812**(2): 162-76.
- Ao Z, Danappa Jayappa K, et al. (2010). "Importin alpha3 interacts with HIV-1 integrase and contributes to HIV-1 nuclear import and replication." *J Virol* **84**(17): 8650-63.
- Ao Z, Fowke KR, et al. (2005). "Contribution of the C-terminal tri-lysine regions of human immunodeficiency virus type 1 integrase for efficient reverse transcription and viral DNA nuclear import." *Retrovirology* **2**(62).
- Ao Z, Huang G, et al. (2007). "Interaction of human immunodeficiency virus type 1 integrase with cellular nuclear import receptor importin 7 and its impact on viral replication." *J Biol Chem* **282**(18): 13456-67.
- Appa RS, Shin CG, et al. (2001). "Role of the nonspecific DNA-binding region and alpha helices within the core domain of retroviral integrase in selecting target DNA sites for integration." *J Biol Chem* **276**(49): 45848-55.
- Arhel NJ, Souquere-Besse S, et al. (2007). "HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore." *EMBO J* **26**(12): 3025-37.
- Ariumi Y, Serhan F, et al. (2006). "The integrase interactor 1 (INI1) proteins facilitate Tat-mediated human immunodeficiency virus type 1 transcription." *Retrovirology* **3**(47).
- Ariumi Y, Turelli P, et al. (2005). "DNA damage sensors ATM, ATR, DNA-PKcs, and PARP-1 are dispensable for human immunodeficiency virus type 1 integration." *J Virol* **79**(5): 2973-8.

## B

- Balada E, Vilardell-Tarrés M, et al. (2010). "Implication of human endogenous retroviruses in the development of autoimmune diseases." *Int Rev Immunol* **29**(4): 351-70.

- Ballandras A, Moreau K, et al. (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**(8) e23032.
- Bao KK, Wang H, et al. (2003). "Functional oligomeric state of avian sarcoma virus integrase." *J Biol Chem* **278**(2): 1323-7.
- Baranova S, Tuzikov FV, et al. (2007). "Small-angle X-ray characterization of the nucleoprotein complexes resulting from DNA-induced oligomerization of HIV-1 integrase." *Nucleic Acids Res* **35**(3): 975-87.
- Basyuk E, Galli T, et al. (2003). "Retroviral genomic RNAs are transported to the plasma membrane by endosomal vesicles." *Dev Cell* **5**(1): 161-74.
- Battle DJ, Kasim M, et al. (2006). "The SMN complex: an assembly machine for RNPs." *Cold Spring Harb Symp Quant Biol* **71**: 313-20.
- Beitzel B and Bushman F (2003). "Construction and analysis of cells lacking the HMGA gene family." *Nucleic Acids Res* **31**(17): 5025-32.
- Belkina AC and Denis GV (2010). "Obesity genes and insulin resistance." *Curr Opin Endocrinol Diabetes Obes* **17**(5): 472-7.
- Bera S, Pandey KK, et al. (2009). "Molecular Interactions between HIV-1 integrase and the two viral DNA ends within the synaptic complex that mediates concerted integration." *J Mol Biol* **389**(1): 183-98.
- Berlitz C and Darlix JL (1995). "An internal ribosomal entry mechanism promotes translation of murine leukemia virus gag polyprotein precursors." *J Virol* **69**(4): 2214-22.
- Biffi A, Bartolomea CC, et al. (2011). "Lentiviral vector common integration sites in preclinical models and a clinical trial reflect a benign integration bias and not oncogenic selection." *Blood* **117**(20): 5332-9.
- Bojja RS, Andrade MD, et al. (2011). "Architecture of a full-length retroviral integrase monomer and dimer, revealed by small angle X-ray scattering and chemical cross-linking." *J Biol Chem* **286**(19): 17047-59.
- Borisenko L (2003). "Avian endogenous retroviruses." *Folia Biol (Praha)* **49**(5): 177-82.
- Bosch S, Arnauld C, et al. (2000). "Study of full-length porcine endogenous retrovirus genomes with envelope gene polymorphism in a specific-pathogen-free Large White swine herd." *J Virol* **74**(18): 8575-81.
- Bowerman B, Brown PO, et al. (1989). "A nucleoprotein complex mediates the integration of retroviral DNA." *Genes Dev* **3**(4): 469-78.
- Brasey A, Lopez-Lastra M, et al. (2003). "The leader of human immunodeficiency virus type 1 genomic RNA harbors an internal ribosome entry segment that is active during the G2/M phase of the cell cycle." *Virologie* **77**(7): 3939-49.
- Brass AL, Dykxhoorn DM, et al. (2008). "Identification of host proteins required for HIV infection through a functional genomic screen." *Science* **319**(5865): 921-6.
- Brin E, Yi J, et al. (2000). "Modeling the late steps in HIV-1 retroviral integrase-catalyzed DNA integration." *J Biol Chem* **275**(50): 39287-95.
- Broomfield S, Hryciw T, et al. (2001). "DNA postreplication repair and mutagenesis in *Saccharomyces cerevisiae*." *Mutat Res* **486**(3): 167-84.
- Bujacz G, Jaskólski M, et al. (1995). "High-resolution structure of the catalytic domain of avian sarcoma virus integrase." *J Mol Biol* **253**(2): 333-46.
- Bushman FD (2007). "Retroviral integration and human gene therapy." *J Clin Invest* **117**(8): 2083-2086.
- Bushman FD, Engelman A, et al. (1993). "Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding." *Proc Natl Acad Sci U S A* **90**(8): 3428-3432.

Busschots K, V. A., De Maeyer M, Rain JC, Emiliani S, Benarous R, Desender L, Debyser Z, Christ F. (2007). "Identification of the LEDGF/p75 binding site in HIV-1 integrase." J Mol Biol **365**(5): 1480-92.

## C

Cai M, Zheng R, et al. (1997). "Solution structure of the N-terminal zinc binding domain of HIV-1 integrase." Nat Struct Biol **4**(7): 567-77.

Carayon K, Leh H, et al. (2010 ). "A cooperative and specific DNA-binding mode of HIV-1 integrase depends on the nature of the metallic cofactor and involves the zinc-containing N-terminal domain." Nucleic Acids Res **38**(11): 3692-708.

Carteau S, Batson SC, et al. (1997). "Human immunodeficiency virus type 1 nucleocapsid protein specifically stimulates Mg<sup>2+</sup>-dependent DNA integration in vitro." J Virol **71**(8): 6225-9.

Carteau S, Gorelick RJ, et al. (1999). "Coupled integration of human immunodeficiency virus type 1 cDNA ends by purified integrase in vitro: stimulation by the viral nucleocapsid protein." J Virol. **73**(8): 6670-9.

Celli GB, Denchi EL, et al. (2006). "Ku70 stimulates fusion of dysfunctional telomeres yet protects chromosome ends from homologous recombination." Nat Cell Biol **8**(8): 885-90.

Cereseto A, Manganaro L, et al. (2005 ). "Acetylation of HIV-1 integrase by p300 regulates viral integration." EMBO J **24**(17): 3070-81.

Chalker DL and Sandmeyer SB (1992). "Ty3 integrates within the region of RNA polymerase III transcription initiation." Genes Dev **6**(1): 117-28.

Chen Z, Yan Y, et al. (2000). "X-ray structure of simian immunodeficiency virus integrase containing the core and C-terminal domain (residues 50-293)--an initial glance of the viral DNA binding platform." J Mol Biol **296**(2): 521-33.

Cherepanov P (2007). "LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro." Nucleic Acids Res. **35**(1): 113-24.

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

Cherepanov P, Maertens G, et al. (2003). "HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells." J Biol Chem **278**(1): 372-81.

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

Christ F, Thys W, et al. (2008). "Transportin-SR2 imports HIV into the nucleus." Curr Biol **18**(16): 1192-202.

Cimarelli A, Sandin S, et al. (2000). "Basic residues in human immunodeficiency virus type 1 nucleocapsid promote virion assembly via interaction with RNA." J Virol **74**(7): 3046-57.

Ciuffi A (2008). "Mechanisms governing lentivirus integration site selection." Curr Gene Ther **8**(6): 419-29.

Ciuffi A, Llano M, et al. (2005). "A role for LEDGF/p75 in targeting HIV DNA integration." Nat Med **11**(12): 1287-9.

Coffin JM, Hughes SH, et al. (1997). "Retroviruses." Cold Spring Harbor Laboratory Press. Bookshelf ID: NBK19361.

Cohen CJ, Lock WM, et al. (2009). "Endogenous retroviral LTRs as promoters for human genes: a critical assessment." Gene **448**(2): 105-14.

- Covey SN (1986). "Amino acid sequence homology in gag region of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus." Nucleic Acids Res **14**(2): 623-33.
- Crise B, Li Y, et al. (2005). "Simian immunodeficiency virus integration preference is similar to that of human immunodeficiency virus type 1." J Virol **79**(19): 12199-204.
- Crittenden LB and Fadly AM (1985). "Responses of chickens lacking or expressing endogenous avian leukosis virus genes to infection with exogenous virus." Poult Sci **64**(3): 454-63.
- Crowley TE, Kaine EM, et al. (2002). "Reproductive cycle regulation of nuclear import, euchromatic localization, and association with components of Pol II mediator of a mammalian double-bromodomain protein." Mol Endocrinol **16**(8): 1727-37.
- Cruceanu M, Urbaneja MA, et al. (2006). "Nucleic acid binding and chaperone properties of HIV-1 Gag and nucleocapsid proteins." Nucleic Acids Res **34**(2): 593-605.

## D

- Daniel R, Katz RA, et al. (2001). "Wortmannin potentiates integrase-mediated killing of lymphocytes and reduces the efficiency of stable transduction by retroviruses." Mol Cell Biol **21**(4): 1164-1172.
- Daniel R and Pomerantz RJ (2005). "ATM: HIV-1's Achilles heel?" Nat Cell Biol **7**: 452-453.
- de Bakker PI, McVean G, et al. (2006). "A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC." Nat Genet **38**(10): 1166-72.
- de Harven E (1974). "Remarks on the ultrastructure of type A, B, and C virus particles." Adv Virus Res **19**: 221-64.
- de Noronha CM, Sherman MP, et al. (2001). "Dynamic disruptions in nuclear envelope architecture and integrity induced by HIV-1 Vpr." Science **294**(5544): 1105-8.
- De Rijck J, Vandekerckhove L, et al. (2006). "Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication." J Virol **80**(23): 11498-509.
- Debysier Z and Christ F (2010). "On the cell biology of HIV integration from basic research to development of novel antiviral drugs." Verh K Acad Geneesk Belg **72**(5-6): 219-37.
- Deffaud C and Darlix JL (2000). "Rous sarcoma virus translation revisited: characterization of an internal ribosome entry segment in the 5' leader of the genomic RNA." J Virol **74**(24): 11581-8.
- Dehart JL, Andersen JL, et al. (2005). "The ataxia telangiectasia-mutated and Rad3-related protein is dispensable for retroviral integration." J Virol **79**(3): 1389-1396.
- Delelis O, Carayon K, et al. (2008). "Integrase and integration: biochemical activities of HIV-1 integrase." Retrovirology **5**(114).
- Delelis O, Parissi V, et al. (2007). "Efficient and specific internal cleavage of a retroviral palindromic DNA sequence by tetrameric HIV-1 integrase." PLoS One **2**(7).
- Denesvre C and Dambrine G (2004). "Les virus leucosiques aviaires ALV-J : origine et maîtrise de l'émergence rétrovirale." Virologie **8**(5): 355-65.
- Denis GV and Green MR (1996). "A novel, mitogen-activated nuclear kinase is related to a Drosophila developmental regulator." Genes Dev **10**(3): 261-71.
- Denis GV, McComb ME, et al. (2006). "Identification of transcription complexes that contain the double bromodomain protein Brd2 and chromatin remodeling machines." J Proteome Res **5**(3): 502-511.

- Denis GV, Nikolajczyk BS, et al. (2010). "An emerging role for bromodomain-containing proteins in chromatin regulation and transcriptional control of adipogenesis." *FEBS Lett* **584**(15).
- Denner J (2008). "Recombinant porcine endogenous retroviruses (PERV-A/C): a new risk for xenotransplantation?" *Arch Virol* **153**(8): 1421-6.
- Denner J (2011). "Infectious risk in xenotransplantation--what post-transplant screening for the human recipient?" *Xenotransplantation* **18**(3): 151-7.
- Deprez E, Barbe S, et al. (2004). "Mechanism of HIV-1 integrase inhibition by styrylquinoline derivatives in vitro." *Mol Pharmacol* **65**(1): 85-98.
- Deprez E, Tauc P, et al. (2000). "Oligomeric states of the HIV-1 integrase as measured by time-resolved fluorescence anisotropy." *Biochemistry* **39**(31): 9275-84.
- Deprez E, Tauc P, et al. (2001). "DNA binding induces dissociation of the multimeric form of HIV-1 integrase: a time-resolved fluorescence anisotropy study." *Proc Natl Acad Sci U S A* **98**(18): 10090-5.
- Derse D, Crise B, et al. (2007). "Human T-cell leukemia virus type 1 integration target sites in the human genome: comparison with those of other retroviruses." *J Virol* **81**(12): 6731-41.
- Desfarges S, San Filippo J, et al. (2006). "Chromosomal integration of LTR-flanked DNA in yeast expressing HIV-1 integrase: down regulation by RAD51." *Nucleic Acids Res* **34**(21): 6215-24.
- Dewannieux M, Ribet D, et al. (2010). "Risks linked to endogenous retroviruses for vaccine production: a general overview." *Biologicals* **38**(3): 366-70.
- Dorsey MJ, Tae HJ, et al. (1995). "B-ATF: a novel human bZIP protein that associates with members of the AP-1 transcription factor family." *Oncogene* **11**(11): 2255-65.
- Drelich M, Wilhelm R, et al. (1992). "Identification of amino acid residues critical for endonuclease and integration activities of HIV-1 IN protein in vitro." *Virology* **188**(2): 459-68.
- Du L, Zhao Y, et al. (2008 ). "D77, one benzoic acid derivative, functions as a novel anti-HIV-1 inhibitor targeting the interaction between integrase and cellular LEDGF/p75." *Biochem Biophys Res Commun.* **375**(1): 139-44.
- Dubé M, Bego MG, et al. (2010). "Modulation of HIV-1-host interaction: role of the Vpu accessory protein." *Retrovirology* **7**(114).
- Dyda F, Hickman AB, et al. (1994). "Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases." *Science* **266**(5193): 1981-6.

## E

- Eichenbaum KD, Rodríguez Y, et al. (2010). "The energetics of the acetylation switch in p53-mediated transcriptional activation." *Proteins* **78**(2): 447-56.
- Eijkelenboom AP, Lutzke RA, et al. (1995). "The DNA-binding domain of HIV-1 integrase has an SH3-like fold." *Nat Struct Biol* **2**(9): 807-10.
- Eijkelenboom AP, Sprangers R, et al. (1999). "Refined solution structure of the C-terminal DNA-binding domain of human immunovirus-1 integrase." *Proteins* **36**(4): 556-64.
- Emiliani S, Mousnier A, Busschots K, Maroun M, Van Maele B, Tempé D, Vandekerckhove L, Moisant F, Ben-Slama L, Witvrouw M, Christ F, Rain JC, Dargemont C, Debyser Z, Benarous R. (2005). "Integrase mutants defective for interaction with LEDGF/p75 are impaired in chromosome tethering and HIV-1 replication." *J Biol Chem.* **280**: 25517-23.

- Engelman, A. (2010). Reverse transcription and integration. Retroviruses: Molecular biology, Genomics and Pathogenesis. R. Kurth and N. Bannert. Norfolk, UK, Caister Academic Press.: 129-159.
- Engelman A, Bushman FD, et al. (1993). "Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex." EMBO J **12**(8): 3269-75.
- Engelman A and Cherepanov P (2008). "The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication." PLoS Pathog **4**(3): 1-9.
- Engelman A and Craigie R (1992). "Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro." J Virol. **66**(11): 6361-9.
- Engelman A and Craigie R (1995). "Efficient magnesium-dependent human immunodeficiency virus type 1 integrase activity." J Virol **69**(9): 5908-5911.
- Engelman A, Mizuuchi K, et al. (1991). "HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer." Cell **67**(6): 1211-21.
- Ericsson T, Oldmixon B, et al. (2001). "Identification of novel porcine endogenous betaretrovirus sequences in miniature swine." J Virol **75**(6): 2765-70.
- Ericsson TA, Takeuchi Y, et al. (2003). "Identification of receptors for pig endogenous retrovirus." Proc Natl Acad Sci U S A **100**(11): 6759-64.
- Espeseth AS, Fishel R, et al. (2011). "siRNA screening of a targeted library of DNA repair factors in HIV infection reveals a role for base excision repair in HIV integration." PLoS One **6**(3): e17612.
- Esposito D and Craigie R (1998). "Sequence specificity of viral end DNA binding by HIV-1 integrase reveals critical regions for protein-DNA interaction." EMBO J **17**(19): 5832-43.

## F

- Falck J, Coates J, et al. (2005). "Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage." Nature **434**(7033): 605-611.
- Fan X, Zhang FH, et al. (2011). "Design of HIV-1 integrase inhibitors targeting the catalytic domain as well as its interaction with LEDGF/p75: A scaffold hopping approach using salicylate and catechol groups." Bioorg Med Chem **19**(16): 4935-52.
- Farnet CM and Bushman FD (1997). "HIV-1 cDNA integration: requirement of HMG I(Y) protein for function of preintegration complexes in vitro." Cell. **88**(4): 483-92.
- Faschinger A, Rouault F, et al. (2008). "Mouse mammary tumor virus integration site selection in human and mouse genomes." J Virol **82**(3): 1360-7.
- Fassati A, Görlich D, et al. (2003). "Nuclear import of HIV-1 intracellular reverse transcription complexes is mediated by importin 7." EMBO J **22**(14): 3675-85.
- Faure A, Calmels C, et al. (2005). "HIV-1 integrase crosslinked oligomers are active in vitro." Nucleic Acids Res **33**(3): 977-86.
- Fischer A, Hacein-Bey-Abina S, et al. (2011). "Gene therapy for primary adaptive immune deficiencies." J Allergy Clin Immunol **127**(6): 1356-9.
- Florence B and Faller DV (2001). "You bet-cha: a novel family of transcriptional regulators." Front Biosci. **6**: 1008-18.
- Frankel AD and Young JA (1998). "HIV-1: fifteen proteins and an RNA." Annu Rev Biochem **67**: 1-25.
- Freed EO (2001). "HIV-1 replication." Somat Cell Mol Genet **26**(1-6): 13-33.

Freed EO, Englund G, et al. (1995). "Role of the basic domain of human immunodeficiency virus type 1 matrix in macrophage infection." J Virol **69**(6): 3949-54.

## G

Gaken J, Tavassoli AM, et al. (1996). "Efficient retroviral infection of mammalian cells is blocked by inhibition of poly(ADP-ribose) polymerase activity." J. Virol **70**: 3992-4000.

Gao K, Gorelick RJ, et al. (2003). "Cofactors for human immunodeficiency virus type 1 cDNA integration in vitro." J Virol **77**(2): 1598-603.

Gerton JL, Ohgi S, et al. (1998). "Effects of mutations in residues near the active site of human immunodeficiency virus type 1 integrase on specific enzyme-substrate interactions." J Virol **72**(6): 5046-55.

Giri MS, Nebozhyn M, et al. (2006). "Microarray data on gene modulation by HIV-1 in immune cells: 2000-2006." J Leukoc Biol **80**(5): 1031-43.

Goldgur Y, Dyda F, et al. (1998). "Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium." Proc Natl Acad Sci U S A. **95**(16): 9150-4.

Goodarzi G, Im GJ, et al. (1995). "Concerted integration of retrovirus-like DNA by human immunodeficiency virus type 1 integrase." J Virol **69**(10): 6090-7.

Gould SJ, Booth AM, et al. (2003). "The Trojan exosome hypothesis." Proc. Natl Acad. Sci. USA **100**(19): 10592-10597.

Greenwald RJ, Tumang JR, et al. (2004). "E mu-BRD2 transgenic mice develop B-cell lymphoma and leukemia." Blood **103**(4): 1475-84.

Greger JG, Katz RA, et al. (2005). "The Cellular Protein Daxx Interacts with Avian Sarcoma Virus Integrase and Viral DNA To Repress Viral Transcription." J Virol. **79**(8): 4610-4618.

Grewe B and Uberla K (2010). "The human immunodeficiency virus type 1 Rev protein: ménage à trois during the early phase of the lentiviral replication cycle." J Gen Virol. **91**(8): 1893-7.

Guiot E, Carayon K, et al. (2006). "Relationship between the oligomeric status of HIV-1 integrase on DNA and enzymatic activity." J Biol Chem **281**(32): 22707-19.

Guntaka RV, Mahy BW, et al. (1975). "Ethidium bromide inhibits appearance of closed circular viral DNA and integration of virus specific DNA in duck cells infected by avian sarcoma virus." Nature **253**(5492): 507-11.

Gyuris A, Donovan DJ, et al. (2009). "The chromatin-targeting protein Brd2 is required for neural tube closure and embryogenesis." Biochim Biophys Acta **1789**(5): 413-421.

## H

Hacein-Bey-Abina S, Von Kalle C, et al. (2003). "LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1." Science **302**(5644): 415-9.

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

Hamamoto S, Nishitsuji H, et al. (2006). "Identification of a novel human immunodeficiency virus type 1 integrase interactor, Gemin2, that facilitates efficient viral cDNA synthesis in vivo." J Virol **80**(12): 5670-7.

- Han Y, Lassen K, et al. (2004). "Resting CD4<sup>+</sup> T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes." *J Virol* **78**(12): 6122-33.
- Hare S, Di Nunzio F, et al. (2009). "Structural basis for functional tetramerization of lentiviral integrase." *PLoS Pathog* **5**(7).
- Hare S, Gupta SS, et al. (2010). "Retroviral intasome assembly and inhibition of DNA strand transfer." *Nature* **464**(7286): 232-6.
- Hare S, Cherepanov P, et al. (2009). "Application of general formulas for the correction of a lattice-translocation defect in crystals of a lentiviral integrase in complex with LEDGF." *Acta Crystallogr D Biol Crystallogr* **65**(Pt 9): 966-73.
- Hare S, Di Nunzio F, et al. (2009). "Structural basis for functional tetramerization of lentiviral integrase." *PLoS Pathog* **5**(7): e1000515.
- Hare S, Gupta SS, et al. (2010). "Retroviral intasome assembly and inhibition of DNA strand transfer." *Nature* **464**(7289): 232-6.
- Harper AL, Sudol M, et al. (2003). "An amino acid in the central catalytic domain of three retroviral integrases that affects target site selection in nonviral DNA." *J Virol* **77**(6): 3838-45.
- Harris D and Engelman A (2000). "Both the structure and DNA binding function of the barrier-to-autointegration factor contribute to reconstitution of HIV type 1 integration in vitro." *J Biol Chem* **275**(50): 39671-7.
- Hartley KO, Gell D, et al. (1995). "DNA-dependent protein kinase catalytic subunit: a relative of phosphatidylinositol 3-kinase and the ataxia telangiectasia gene product." *Cell* **82**(5): 849-56.
- Hasan S, Stucki M, et al. (2001). "Regulation of human flap endonuclease-1 activity by acetylation through the transcriptional coactivator p300." *Mol Cell Bio* **7**(6): 1221-31.
- Hayouka Z, Rosenbluh J, et al. (2007). "Inhibiting HIV-1 integrase by shifting its oligomerization equilibrium." *Proc Natl Acad Sci U S A* **104**(20): 8316-21.
- Hazuda DJ (2010). "Resistance to inhibitors of the human immunodeficiency virus type 1 integration." *Braz J Infect Dis* **14**(5): 513-8.
- Hazuda DJ, Felock P, et al. (2000). "Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells." *Science*. **287**(5453): 646-50.
- Hearps AC, Wagstaff KM, et al. (2008). "The N-terminal basic domain of the HIV-1 matrix protein does not contain a conventional nuclear localization sequence but is required for DNA binding and protein self-association." *Biochemistry* **47**(7): 2199-210.
- Hemenway CS, de Erkenez AC, et al. (2001). "The polycomb protein MPC3 interacts with AF9, an MLL fusion partner in t(9;11)(p22;q23) acute leukemias." *Oncogene* **20**(29): 3798-805.
- Heuer TS, and Brown PO (1997). "Mapping features of HIV-1 integrase near selected sites on viral and target DNA molecules in an active enzyme-DNA complex by photo-cross-linking." *Biochemistry* **36**(35): 10655-65.
- Heuer TS, and Brown PO (1998). "Photo-cross-linking studies suggest a model for the architecture of an active human immunodeficiency virus type 1 integrase-DNA complex." *Biochemistry* **37**(19): 6667-78.
- Hindmarsh P and Leis J (1999). "Retroviral DNA integration." *Microbiol Mol Biol Rev* **63**(4): 836-43.
- Hindmarsh P, Ridky T, et al. (1999). "HMG protein family members stimulate human immunodeficiency virus type 1 and avian sarcoma virus concerted DNA integration in vitro." *J Virol* **73**(4): 2994-3003.

- Hombrouck A, De Rijck J, et al. (2007). "Virus evolution reveals an exclusive role for LEDGF/p75 in chromosomal tethering of HIV." *PLoS Pathog* **3**(3).
- Huang L, Xu GL, et al. (2008). "Daxx interacts with HIV-1 integrase and inhibits lentiviral gene expression." *Biochem Biophys Res Commun* **373**(2): 241-5.
- Huda A, Polavarapu N, et al. (2008). "Endogenous retroviruses of the chicken genome." **3** **9**.

## I

- Ikeda T, Nishitsuji H, et al. (2004 ). "Evaluation of the functional involvement of human immunodeficiency virus type 1 integrase in nuclear import of viral cDNA during acute infection." *J Virol* **78**(21): 11563-73.
- Inagaki A, Sleddens-Linkels E, et al. (2011). "Human RAD18 Interacts with Ubiquitylated Chromatin Components and Facilitates RAD9 Recruitment to DNA Double Strand Breaks." *PLoS One* **6**(8).
- Inayoshi Y, Okino Y, et al. (2010). "Transcription factor YY1 interacts with retroviral integrases and facilitates integration of moloney murine leukemia virus cDNA into the host chromosomes." *J Virol* **84**(16): 8250-61.

## J

- Jang MK, Kwon D, et al. (2009). "Papillomavirus E2 proteins and the host BRD4 protein associate with transcriptionally active cellular chromatin." *J Virol* **83**(6): 2592-600.
- Jayappa KD, Ao Z, et al. (2011). "Identification of Critical Motifs within HIV-1 Integrase Required for Importin  $\alpha$ 3 Interaction and Viral cDNA Nuclear Import." *J Mol Biol* **410**: 847-62.
- Jenkins TM, Espesito D, Engelman A, Craigie R. (1997). "Critical contacts between HIV-1 integrase and viral DNA identified by structure-based analysis and photo-crosslinking." *EMBO J* **16**(22): 6849-59.
- Jenkins TM, Engelman A, et al. (1996). "A soluble active mutant of HIV-1 integrase: involvement of both the core and carboxyl-terminal domains in multimerization." *J Biol Chem* **271**(13): 7712-8.
- Jenkins TM, Esposito D, et al. (1997). "Critical contacts between HIV-1 integrase and viral DNA identified by structure-based analysis and photo-crosslinking." *Embo J* **16**(22): 6849-59.
- Johnson MS, McClure MA, et al. (1986). "Computer analysis of retroviral pol genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes." *Proc Natl Acad Sci U S A* **83**(20): 7648-52.
- Johri MK, Mishra R, et al. (2011). "Tits and bits of HIV Tat protein." *Expert Opin Biol Ther* **11**(3): 269-83.
- Jonsson CB, Donzella G A, et al. (1996). "Functional domains of Moloney murine leukemia virus integrase defined by mutation and complementation analysis." *J Virol* **70**: 4585-97.
- Jonsson CB, Donzella GA, et al. (1993). "Characterization of the forward and reverse integration reactions of the Moloney murine leukemia virus integrase protein purified from Escherichia coli." *J Biol Chem* **268**(2): 1462-9.
- Jordan A, Bisgrove D, et al. (2003). "HIV reproducibly establishes a latent infection after acute infection of T cells in vitro." *EMBO J* **22**(8): 1868-77.

Jordan A, Defechereux P, et al. (2001). "The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation." EMBO J **20**(7): 1726-38.

## K

Kalpana GV, Marmon S, et al. (1994). "Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5." Science **266**(5193): 2002-6.

Kameoka M, Nukuzuma S, et al. (2005). "Poly(ADP-ribose)polymerase-1 is required for integration of the human immunodeficiency virus type 1 genome near centromeric alphoid DNA in human and murine cells." Biochem Biophys Res Commun. **334**(2): 412-7.

Kang Y, Moressi CJ, et al. (2006). "Integration site choice of a feline immunodeficiency virus vector." J Virol **80**(17): 8820-3.

Kanno T, Kanno Y, et al. (2004). "Selective recognition of acetylated histones by bromodomain proteins visualized in living cells." Mol Cell **13**(1): 33-43.

Katz RA, Greger JG, et al. (2002). "Transduction of interphase cells by avian sarcoma virus." J Virol **76**(11): 5422-34.

Katzman M, and Katz RA (1999). "Substrate recognition by retroviral integrases." Adv Virus Res **52**: 371-95.

Katzman M, and Sudol M (1995). "Mapping domains of retroviral integrase responsible for viral DNA specificity and target site selection by analysis of chimeras between human immunodeficiency virus type 1 and visna virus integrases." J Virol **69**(9): 5687-5696.

Katzman M, and Sudol M (1998). "Mapping viral DNA specificity to the central region of integrase by using functional human immunodeficiency virus type 1/visna virus chimeric proteins." J Virol **72**(3): 1744-1753.

Katzourakis A, Tristem M, et al. (2007). "Discovery and analysis of the first endogenous lentivirus." Proc Natl Acad Sci U S A **104**(15): 6261-5.

Kessl JJ, Li M, Ignatov M, et al. (2011). "FRET analysis reveals distinct conformations of IN tetramers in the presence of viral DNA or LEDGF/p75." Nucleic Acids Res.

Khan E, Mack JP, et al. (1991). "Retroviral integrase domains: DNA binding and the recognition of LTR sequences." Nucleic Acids Res. **19**(4): 851-60.

Kobayashi J, Tauchi H, et al. (2002). "NBS1 localizes to gamma-H2AX foci through interaction with the FHA/BRCT domain." Curr Biol. **12**(21): 1846-1851.

Konig R, Zhou Y, et al. (2008). "Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication." Cell Cycle **135**(1): 49-60.

Krishnan L, Li X, et al. (2010). "Structure-based modeling of the functional HIV-1 intasome and its inhibition." Proc Natl Acad Sci U S A **107**(36): 15910-5.

Krishnan L, Matreyek KA, et al. (2010). "The requirement for cellular transportin 3 (TNPO3 or TRN-SR2) during infection maps to human immunodeficiency virus type 1 capsid and not integrase." J Virol **84**(1): 397-406.

Kuhnlein U, Fairfull RW, et al. (1993). "Synergism between the endogenous viral loci ev6 and ev9 in inducing immunological tolerance to avian leukosis virus." Br Poult Sci **34**(1): 93-104.

Kukolj G, Katz RA, et al. (1998). "Characterization of the nuclear localization signal in the avian sarcoma virus integrase." Gene **223**(1-2): 157-63.

Kulkarni A, Pavithra L, et al. (2004). "HIV-1 integration sites are flanked by potential MARs that alone can act as promoters." Biochem Biophys Res Commun **322**(2): 672-7.

## L

- Lachuer J, Ronfort C, et al. (1996). "Characterization of a cDNA encoding an alpha thyroid hormone receptor in muscovy duckling." *Poult Sci* **75**(12): 1531-5.
- Lander ES, Linton LM, et al. (2001). "Initial sequencing and analysis of the human genome." *Nature* **409**(6822): 860-921.
- Lau A, Swinbank KM, et al. (2005). "Suppression of HIV-1 infection by a small molecule inhibitor of the ATM kinase." *Nat Cell Biol*. **7**: 493-500.
- Le Tissier P, Stoye JP, et al. (1997). "Two sets of human-tropic pig retrovirus." *Nature* **389**(6652): 681-2.
- Lee D, Kim NY, et al. (2008). "Transmissible infection of human 293T cells with porcine endogenous retroviruses subgroup a from NIH-miniature pig." *Transplant Proc* **40**(10): 3742-5.
- Lee KK, Haraguchi T, et al. (2001). "Distinct functional domains in emerin bind lamin A and DNA-bridging protein BAF." *J Cell Sci* **114**(24): 4567-73.
- Lee MS and Craigie R (1994). "Protection of retroviral DNA from autointegration: involvement of a cellular factor." *Proc Natl Acad Sci U S A* **91**(21): 9823-7.
- Lee MS and Craigie R (1998). "A previously unidentified host protein protects retroviral DNA from autointegration." *Proc Natl Acad Sci U S A* **95**(4): 1528-1533.
- Lees-Miller SP and Meek K (2003). "Repair of DNA double strand breaks by non-homologous end joining." *Biochimie* **85**(11): 1161-73.
- Leh H, Brodin P, et al. (2000). "Determinants of Mg<sup>2+</sup>-dependent activities of recombinant human immunodeficiency virus type 1 integrase." *Biochemistry* **39**(31): 9285-94.
- LeRoy G, Rickards B, et al. (2008). "The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription." *Mol Cell Biol*. **30**(1): 51-60.
- Lesbats P, Botbol Y, et al. (2011). "Functional coupling between HIV-1 integrase and the SWI/SNF chromatin remodeling complex for efficient in vitro integration into stable nucleosomes." *PLoS Pathog* **7**(2).
- Levin A, Armon-Omer A, et al. (2009). "Inhibition of HIV-1 integrase nuclear import and replication by a peptide bearing integrase putative nuclear localization signal." *Retrovirology* **6**(112).
- Lewinski MK, Bisgrove D, et al. (2005). "Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription." *J Virol* **79**(11): 6610-9.
- Lewinski MK, Yamashita M, et al. (2006). "Retroviral DNA integration: viral and cellular determinants of target-site selection." *PLoS Pathog* **2**(6): 611-622.
- Li L, Li HS, et al. (2005). "Roles of HIV-1 auxiliary proteins in viral pathogenesis and host-pathogen interaction." *Cell Res* **15**(11-12): 923-34.
- Li M, Mizuuchi M, et al. (2006). "Retroviral DNA integration: reaction pathway and critical intermediates." *Embo J* **25**(6): 1295-304.
- Li X, Krishnan L, et al. (2011). "Structural biology of retroviral DNA integration." *Virology* **411**(2): 194-205.
- Lin CW and Engelman A (2003). "The barrier-to-autointegration factor is a component of functional human immunodeficiency virus type 1 preintegration complexes." *J Virol*. **77**(8): 5030-6.
- Liu XH, Xu W, et al. (2011). "The host range of gammaretroviruses and gammaretroviral vectors includes post-mitotic neural cells." *PLoS One* **6**(3).
- Liu Y, Kao HI, et al. (2004). "Flap endonuclease 1: a central component of DNA metabolism." *Annu Rev Biochem* **73**: 589-615.

- Llano M, Vanegas M, et al. (2004). "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**(17): 9524-37.
- Lloyd AG, Tateishi S, et al. (2006). "Effect of DNA repair protein Rad18 on viral infection." *PLoS Pathog* **2**(5).
- Lodi PJ, Ernst JA, et al. (1995). "Solution structure of the DNA binding domain of HIV-1 integrase." *Biochemistry* **34**(31): 9826-33.
- Lu R, Limon A, et al. (2004). "Class II integrase mutants with changes in putative nuclear localization signals are primarily blocked at a postnuclear entry step of human immunodeficiency virus type 1 replication." *J Virol* **78**(23): 12735-12746.
- Lukas C, Falck J, et al. (2003). "Distinct spatiotemporal dynamics of mammalian checkpoint regulators induced by DNA damage." *Nat Cell Biol* **5**(3): 255-260.
- Lutzke RA and Plasterk RH (1998). "Structure-based mutational analysis of the C-terminal DNA-binding domain of human immunodeficiency virus type 1 integrase: critical residues for protein oligomerization and DNA binding." *J Virol* **72**(6): 4841-8.
- Lutzke RA, Vink C, et al. (1994). "Characterization of the minimal DNA-binding domain of the HIV integrase protein." *Nucleic Acids Res.* **22**(20): 4125-31.

## M

- Maertens GN, Hare S, et al. (2010). "The mechanism of retroviral integration from X-ray structures of its key intermediates." *Nature* **468**(7321): 326-9.
- Mannioui A, Nelson E, et al. (2005). "Human immunodeficiency virus type 1 KK26-27 matrix mutants display impaired infectivity, circularization and integration but not nuclear import." *Virology* **339**(1): 21-30.
- Mantovani F, Drost J, et al. (2010). "Gene regulation and tumor suppression by the bromodomain-containing protein BRD7." *Cell Cycle* **9**(14): 2777-81.
- Maroun M, Delelis O, et al. (2006). "Inhibition of early steps of HIV-1 replication by SNF5/Ini1." *J Biol Chem* **281**(32): 22736-43.
- Marsden MD and Zack JA (2007). "Human immunodeficiency virus bearing a disrupted central DNA flap is pathogenic in vivo." *J Virol* **81**(11): 6146-50.
- Martin GS (2004). "The road to Src." *Oncogene* **23**(48): 7910-7.
- Mattsson K, Kiss C, et al. (2002). "Latent nuclear antigen of Kaposi's sarcoma herpesvirus/human herpesvirus-8 induces and relocates RING3 to nuclear heterochromatin regions." *J Gen Virol* **83**(1): 179-88.
- Maxfield LF, Fraize CD, et al. (2005). "Relationship between retroviral DNA-integration-site selection and host cell transcription." *Proc Natl Acad Sci U S A* **102**(5): 1436-41.
- Mayer J and Meese E (2005). "Human endogenous retroviruses in the primate lineage and their influence on host genomes." *Cytogenet Genome Res* **110**(1-4): 448-56.
- McKee CJ, Kessler JJ, et al. (2008). "Dynamic modulation of HIV-1 integrase structure and function by cellular lens epithelium-derived growth factor (LEDGF) protein." *J Biol Chem* **283**: 31802-31812.
- Meehan, A. M. and E. M. Poeschla (2010). "Chromatin tethering and retroviral integration: recent discoveries and parallels with DNA viruses." *Biochim Biophys Acta* **1799**(3-4): 182-91.
- Mi S, Lee X, et al. (2000). "Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis." *Nature* **403**(6771): 785-9.
- Michel F, Crucifix C, et al. (2009). "Structural basis for HIV-1 DNA integration in the human genome, role of the LEDGF/P75 cofactor." *EMBO J* **28**(7): 980-91.

- Miller MD, Farnet CM, et al. (1997). "Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition." *J Virol* **71**(7): 5382-90.
- Mitchell RS, Beitzel BF, et al. (2004). "Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences." *PLoS Biol* **2**(8): 1127-1137.
- Moalic Y, Blanchard Y, et al. (2006). "Porcine endogenous retrovirus integration sites in the human genome: features in common with those of murine leukemia virus." *J Virol* **80**(22): 10980-8.
- Moalic Y, Felix H, et al. (2011). "Risque rétroviral endogène porcin (PERV) en xénotransplantation." *Virologie* **15**(1): 63-72.
- Moalic Y, Félix H, et al. (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**(4): 1920-9.
- Moalic Y, Felix H, et al. (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**(4): 1920-9.
- Montes de Oca R, Lee KK, et al. (2005). "Binding of barrier to autointegration factor (BAF) to histone H3 and selected linker histones including H1.1." *J Biol Chem* **280**(51): 42252-62.
- Moreau K, Faure C, et al. (2004). "Mutational analyses of the core domain of Avian Leukemia and Sarcoma Viruses integrase: critical residues for concerted integration and multimerization." *Virology* **318**(2): 566-81.
- Moreau K, Faure C, et al. (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**(22): 4426-38.
- Moreau K, Charmetant J, et al. (2009). "Avian sarcoma and leukemia virus (ASLV) integration in vitro: mutation or deletion of integrase (IN) recognition sequences does not prevent but only reduces the efficiency and accuracy of DNA integration." *Virology* **392**(1): 94-102.
- Moreau K, C. J., Gallay K, Faure C, Verdier G, Ronfort C. (2009). "Avian sarcoma and leukemia virus (ASLV) integration in vitro: mutation or deletion of integrase (IN) recognition sequences does not prevent but only reduces the efficiency and accuracy of DNA integration." *Virology* **392**: 94-102.
- Moreau, K., C. Faure, et al. (2002). "Analysis of conserved and non-conserved amino acids critical for ALSV (Avian leukemia and sarcoma viruses) integrase functions in vitro." *Arch Virol* **147**(9): 1761-78.
- Moreau K, Faure C, et al. (2004). "Mutational analyses of the core domain of Avian Leukemia and Sarcoma Viruses integrase: critical residues for concerted integration and multimerization." *Virology* **318**(2): 566-81.
- Moreau K, Faure C, et al. (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**(22): 4426-38.
- Mousnier A, Kubat N, et al. (2007). "von Hippel Lindau binding protein 1-mediated degradation of integrase affects HIV-1 gene expression at a postintegration step." *Proc Natl Acad Sci U S A* **104**(34): 13615-20.
- Mulder LC, Chakrabarti LA, et al. (2002). "Interaction of HIV-1 integrase with DNA repair protein hRad18." *J Biol Chem* **277**(30): 27489-93.
- Mumm SR and Grandgenett DP (1991). "Defining nucleic acid-binding properties of avian retrovirus integrase by deletion analysis." *J Virol* **65**(3): 1160-1167.

- Muriaux D, Costes KS, et al. (2004). "Role of murine leukemia virus nucleocapsid protein in virus assembly." J Virol **78**: 12378-12385.
- Muriaux D and Darlix JL (2010). "Properties and functions of the nucleocapsid protein in virus assembly." RNA Biol **7**(6): 744-53.

## N

- Nakamura Y, Umehara T, et al. (2007). "Crystal structure of the human BRD2 bromodomain: insights into dimerization and recognition of acetylated histone H4." J Biol Chem **282**(6): 4193-4201.
- Narezkina A, Taganov KD, et al. (2004). "Genome-wide analyses of avian sarcoma virus integration sites." J Virol **78**(21): 11656-63.
- Nethe M, Berkhout B, et al. (2005). "Retroviral superinfection resistance." Retrovirology **2**(52).
- Nguyen BY, Isaacs RD, et al. (2011). "Raltegravir: the first HIV-1 integrase strand transfer inhibitor in the HIV armamentarium." Ann N Y Acad Sci **1222**: 83-9.
- Nishitsuji H, Hayashi T, et al. (2009). "Augmentation of reverse transcription by integrase through an interaction with host factor, SIP1/Gemin2 Is critical for HIV-1 infection." PLoS One **4**(11).
- Nowrouzi A, Dittrich M, et al. (2006). "Genome-wide mapping of foamy virus vector integrations into a human cell line." J Gen Virol **87**: 1339-47.

## O

- Oldmixon BA, Wood JC, et al. (2002). "Porcine endogenous retrovirus transmission characteristics of an inbred herd of miniature swine." J Virol **76**(6): 3045-8.
- Ottinger M, Christalla T, et al. (2006). "Kaposi's 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." J Virol **80**(21): 10772-86.

## P

- Pandey KK, Bera S, et al. (2011). "The HIV-1 integrase monomer induces a specific interaction with LTR DNA for concerted integration." Biochemistry **50**(45): 9788-96.
- Pandey KK, Sinha S, et al. (2007). "Transcriptional coactivator LEDGF/p75 modulates human immunodeficiency virus type 1 integrase-mediated concerted integration." J Virol **81**(8): 3969-79.
- Parent M, Yung TM, et al. (2005). "Poly(ADP-ribose) polymerase-1 is a negative regulator of HIV-1 transcription through competitive binding to TAR RNA with Tat-positive transcription elongation factor b (p-TEFb) complex." J Biol Chem **280**(1): 448-57.
- Parissi V, Calmels C, et al. (2001). "Functional interactions of human immunodeficiency virus type 1 integrase with human and yeast HSP60." J Virol **75**(23): 11344-53.
- Patience C, Switzer WM, et al. (2001). "Multiple groups of novel retroviral genomes in pigs and related species." J Virol **75**(6): 2771-5.
- Patience C, Takeuchi Y, et al. (1997). "Infection of human cells by an endogenous retrovirus of pigs." Nat Med **3**(3): 282-286.
- Payne LN (1998). "HPRS-103 a retrovirus strikes back. The emergence of Subgroup J Avian Leukosis Virus." Avian Pathol **27**: 536-545.

- Peytavi R, Hong SS, et al. (1999). "HEED, the product of the human homolog of the murine eed gene, binds to the matrix protein of HIV-1." J Biol Chem **274**(3): 1635-45.
- Platt GM, Simpson GR, et al. (1999). "Latent nuclear antigen of Kaposi's sarcoma-associated herpesvirus interacts with RING3, a homolog of the Drosophila female sterile homeotic (fsh) gene." J Virol, **73**(12): 9789-95.
- Preuss T, Fischer N, et al. (2006). "Isolation and characterization of an infectious replication-competent molecular clone of ecotropic porcine endogenous retrovirus class C." J Virol **80**(20): 10258-61.
- Prizan-Ravid A, Elis E, et al. (2010). "The Gag Cleavage Product, p12, is a Functional Constituent of the Murine Leukemia Virus Pre-Integration Complex." PLoS Pathog **6**(11).
- Pruss D, Bushman FD, et al. (1994). "Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core." Proc. Natl Acad. Sci. USA **91**(13): 5913-5917.
- Pryciak PM, Sil A, et al. (1992). "Retroviral integration into minichromosomes in vitro." EMBO J **11**(1): 291-303.
- Pryciak PM and Varmus HE (1992). "Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection." Cell **69**(5): 769-780.

## Q

- Qasim W, Gaspar HB, et al. (2009). "Progress and prospects: gene therapy for inherited immunodeficiencies." Gene Ther **16**(11): 1285-91.
- Qiu C, Sawada K, et al. (2002). "The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds." Nat Struct Biol **9**(3): 217-24.

## R

- Raghavendra NK and Engelman A (2007). "LEDGF/p75 interferes with the formation of synaptic nucleoprotein complexes that catalyze full-site HIV-1 DNA integration in vitro: implications for the mechanism of viral cDNA integration." Virology **360**(1): 1-5.
- Rain JC, Cribier A, et al. (2009). "Yeast two-hybrid detection of integrase-host factor interactions." Methods, **47**(4): 291-7.
- Rakotobe D, Tardy JC, et al. (2007). "Human Polycomb group EED protein negatively affects HIV-1 assembly and release." Retrovirology **4**(37).
- Rakotobe D, Violot S, et al. (2008). "Mapping of immunogenic and protein-interacting regions at the surface of the seven-bladed beta-propeller domain of the HIV-1 cellular interactor EED." Virology, **5**(32).
- Reil H, Bukovsky AA, et al. (1998). "Efficient HIV-1 replication can occur in the absence of the viral matrix protein." EMBO J **17**(9): 2699-708.
- Ren G, Gao K, et al. (2007). "Single-particle image reconstruction of a tetramer of HIV integrase bound to DNA." J Mol Biol **366**(1): 286-294.
- Ribet D and Heidmann T (2010). "Formation et évolution des rétrovirus endogènes." Virologie **14**(2).
- Rivière L, Darlix JL, et al. (2010). "Analysis of the viral elements required in the nuclear import of HIV-1 DNA." J Virol **84**(2): 729-39.
- Roe T, Reynolds TC, et al. (1993). "Integration of murine leukemia virus DNA depends on mitosis." EMBO J **12**(5): 2099-108.

- Roelofs H, van Gurp RJ, et al. (1998). "Detection of human endogenous retrovirus type K-specific transcripts in testicular parenchyma and testicular germ cell tumors of adolescents and adults: clinical and biological implications." Am J Pathol **153**(4): 1277-82.
- Ronfort C, De Breyne S, et al. (2004). "Characterization of two distinct RNA domains that regulate translation of the Drosophila gypsy retroelement." RNA **10**(3): 504-15.

## S

- Sandmeyer S (2003). "Integration by design." Proc Natl Acad Sci U S A **100**(10): 5586-8.
- Sbodio JI and Chi NW (2002). "Identification of a tankyrase-binding motif shared by IRAP, TAB182, and human TRF1 but not mouse TRF1. NuMA contains this RXXPDG motif and is a novel tankyrase partner." J Biol Chem **277**(35): 31887-92.
- Schneider MK and Seebach JD (2011). "Xenotransplantation literature update, January-February 2011." Xenotransplantation **18**(2): 147-50.
- Schröder AR, Shinn P, et al. (2002). "HIV-1 integration in the human genome favors active genes and local hotspots." Cell **110**(4): 521-9.
- Seelamgari A, Maddukuri A, et al. (2004). "Role of viral regulatory and accessory proteins in HIV-1 replication." Front Biosci **9**: 2388-413.
- Seimiya H and Smith S (2002). "The telomeric poly(ADP-ribose) polymerase, tankyrase 1, contains multiple binding sites for telomeric repeat binding factor 1 (TRF1) and a novel acceptor, 182-kDa tankyrase-binding protein (TAB182)." J Biol Chem **277**(16): 14116-26.
- Shan L, Yang HC, et al. (2011). "Influence of host gene transcription level and orientation on HIV-1 latency in a primary-cell model." J Virol **85**(11): 5384-93.
- Shang E, Wang X, et al. (2009). "Double bromodomain-containing gene Brd2 is essential for embryonic development in mouse." Dev Dyn **238**(4): 908-917.
- Shibagaki Y and Chow SA (1997). "Central core domain of retroviral integrase is responsible for target site selection." J Biol Chem **272**(49): 8361-9.
- Shumaker DK, Lee KK, et al. (2001). "LAP2 binds to BAF.DNA complexes: requirement for the LEM domain and modulation by variable regions." EMBO J **20**(7): 1754-64.
- Shun MC, Raghavendra NK, et al. (2007). "LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration." Genes Dev **21**(14): 1767-78.
- Silvers RM, Smith JA, et al. (2010). "Modification of integration site preferences of an HIV-1-based vector by expression of a novel synthetic protein." Hum Gene Ther **21**(3): 337-49.
- Sinha A, Faller DV, et al. (2005). "Bromodomain analysis of Brd2-dependent transcriptional activation of cyclin A." Biochem J **387**(1): 257-269.
- Sinha S and Grandgenett DP (2005). "Recombinant human immunodeficiency virus type 1 integrase exhibits a capacity for full-site integration in vitro that is comparable to that of purified preintegration complexes from virus-infected cells." J Virol **79**(13): 8208-16.
- Sinha S, Pursley MH, et al. (2002). "Efficient concerted integration by recombinant human immunodeficiency virus type 1 integrase without cellular or viral cofactors." J Virol **76**(7): 3105-13.
- Skalka AM and Katz RA (2005). "Retroviral DNA integration and the DNA damage response." Cell Death Differ **12**(1): 971-8.

- Skoko D, Li M, et al. (2009). "Barrier-to-autointegration factor (BAF) condenses DNA by looping." *Proc Natl Acad Sci U S A* **106**(39): 16610-5.
- Smith JA, Wang FX, et al. (2008). "Evidence that the Nijmegen breakage syndrome protein, an early sensor of double-strand DNA breaks (DSB), is involved in HIV-1 post-integration repair by recruiting the ataxia telangiectasia-mutated kinase in a process similar to, but distinct from, cellular DSB repair." *Virology* **5**(11).
- Smith JA and Daniel R (2006). "Following the path of the virus: the exploitation of host DNA repair mechanisms by retroviruses." *ACS Chem Biol* **1**(4): 217-26.
- Srinivasan RS, de Erkenez AC, et al. (2003). "The mixed lineage leukemia fusion partner AF9 binds specific isoforms of the BCL-6 corepressor." *Oncogene* **22**(22): 3395-406.
- Stec I, Nagl SB, et al. (2000). "The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation?" *FEBS Lett* **473**(1): 1-5.
- Studamire B and Goff SP (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 (2002). "Regulatory mechanisms by which barrier-to-autointegration factor blocks autointegration and stimulates intermolecular integration of Moloney murine leukemia virus preintegration complexes." *J Virol* **76**(23): 12376-80.
- Suzuki Y, Yang H, et al. (2004). "LAP2alpha and BAF collaborate to organize the Moloney murine leukemia virus preintegration complex." *EMBO J* **23**(23): 4670-8.
- Switzer WM, Michler RE, et al. (2001). "Lack of cross-species transmission of porcine endogenous retrovirus infection to nonhuman primate recipients of porcine cells, tissues, or organs." *Transplantation* **71**(7): 959-65.

## T

- Taganov KD, Cuesta I, et al. (2004). "Integrase-specific enhancement and suppression of retroviral DNA integration by compacted chromatin structure in vitro." *J Virol* **78**(11): 5848-55.
- Takeda A and Matano T (2007). "Inhibition of infectious murine leukemia virus production by Fv-4 env gene products exerting dominant negative effect on viral envelope glycoprotein." *Microbes Infect* **9**(14): 1590-6.
- Takeuchi Y, Patience C, et al. (1998). "Host range and interference studies of three classes of pig endogenous retrovirus." *J Virol* **72**(12): 9986-91.
- Tamkun JW, Deuring R, et al. (1992). "brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2." *Cell Mol Life Sci* **68**(3): 561-572.
- Tarlinton RE, Meers J, et al. (2006). "Retroviral invasion of the koala genome." *Nature* **442**(7098): 79-81.
- Tauchi H, Kobayashi J, et al. (2001). "The forkhead-associated domain of NBS1 is essential for nuclear foci formation after irradiation but not essential for hRAD50[hMRE11]NBS1 complex DNA repair activity." *J Biol Chem* **276**(1): 12-15.
- Topper M, Luo Y, et al. (2007). "Posttranslational acetylation of the human immunodeficiency virus type 1 integrase carboxyl-terminal domain is dispensable for viral replication." *J Virol* **81**(6): 3012-7.
- Torne-Celer, C., K. Moreau, et al. (2008). "Gene transfer into mammalian cells using a self-deleting avian leukemia and sarcoma virus-based vector." *Intervirology* **51**(6): 447-52.

- Tréand C, du Chéné I, et al. (2006). "Requirement for SWI/SNF chromatin-remodeling complex in Tat-mediated activation of the HIV-1 promoter." *EMBO J* **25**(8): 1690-9.
- Trobridge GD, Miller DG, et al. (2006). "Foamy virus vector integration sites in normal human cells." *Proc Natl Acad Sci U S A* **103**(5): 1498-503.
- Tsuchiya H, Iseda T, et al. (1996). "Identification of a novel protein (VBP-1) binding to the von Hippel-Lindau (VHL) tumor suppressor gene product." *Cancer Res* **56**(13): 2881-5.
- Tsuruno C, Ohe K, et al. (2011). "HMGA1a is involved in specific splice site regulation of human immunodeficiency virus type 1." *Biochem Biophys Res Commun* **406**(4): 512-7.
- Tsurutani N, Kubo M, Maeda Y, Ohashi T, Yamamoto N, Kannagi M, Masuda T. (2000). "Identification of critical amino acid residues in human immunodeficiency virus type 1 IN required for efficient proviral DNA formation at steps prior to integration in dividing and nondividing cells." *J Virol* **74**(10): 4795-806.
- Turelli P, Doucas V, et al. (2001). "Cytoplasmic recruitment of INI1 and PML on incoming HIV preintegration complexes: interference with early steps of viral replication." *Mol Cell* **7**(6): 1245-54.
- Turlure F, Devroe E, et al. (2004). "Human cell proteins and human immunodeficiency virus DNA integration." *Front Biosci* **1**(9): 3187-208.
- Turlure F, Maertens G, et al. (2006). "A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo." *Nucleic Acids Res* **34**(5): 1653-65.
- Tyagi S, Ochem A, et al. (2011). "DNA-dependent protein kinase interacts functionally with the RNA polymerase II complex recruited at the human immunodeficiency virus (HIV) long terminal repeat and plays an important role in HIV gene expression." *J Gen Virol* **92**(7): 1710-20.

## U

- Umehara T, Nakamura Y, et al. (2010). "Structural basis for acetylated histone H4 recognition by the human BRD2 bromodomain." *J Biol Chem* **285**(10): 7610-7618.

## V

- Vainberg IE, Lewis SA, et al. (1998). "Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin." *Cell* **93**(5): 863-73.
- van den Ent FM, Vos A, et al. (1999). "Dissecting the role of the N-terminal domain of human immunodeficiency virus integrase by trans-complementation analysis." *J Virol* **73**(4): 3176-83.
- van Gent DC, Vink C, et al. (1993). "Complementation between HIV integrase proteins mutated in different domains." *EMBO J* **12**(8): 3261-7.
- van Lohuizen M, Tijms M, et al. (1998). "Interaction of mouse polycomb-group (Pc-G) proteins Enx1 and Enx2 with Eed: indication for separate Pc-G complexes." *Mol Cell Bio* **18**(6): 3572-9.
- Varmus HE, Guntaka RV, et al. (1975). "Synthesis, structure and function of avian sarcoma virus-specific DNA in permissive and nonpermissive cells." *Cold Spring Harb Symp Quant Biol* **39** (2): 987-96.
- Venugopal K (1999). "Avian leukosis virus subgroup J: a rapidly evolving group of oncogenic retroviruses." *Res Vet Sci* **67**(2): 113-9.

- Vera J, Parissi V, Garcia A, Zuniga R, Androela ML, Caumont-Sacros A, Tarrago-Litvak L and Leon O (2005). "Yeast system as a model to study moloney murine leukemia virus integrase : expression, mutagenesis and search for eukaryotic partners." J. Gen. Virol **86**: 2481-2488.
- Viejo-Borbolla A, Ottinger M, et al. (2005). "Brd2/RING3 interacts with a chromatin-binding domain in the Kaposi's Sarcoma-associated herpesvirus latency-associated nuclear antigen 1 (LANA-1) that is required for multiple functions of LANA-1." J Virol **79**(21): 13618-29.
- Villanueva RA, Jonsson CB, et al. (2003). "Differential multimerization of Moloney murine leukemia virus integrase purified under nondenaturing conditions." Virology **316**(1): 146-60.
- Vink C, Yeheskiely E, et al. (1991). "Site-specific hydrolysis and alcoholysis of human immunodeficiency virus DNA termini mediated by the viral integrase protein." Nucleic Acids Res. **19**(24): 6691-8.
- Violot S, Hong SS, et al. (2003). "The human polycomb group EED protein interacts with the integrase of human immunodeficiency virus type 1." J Virol **77**(23): 12507-22.
- Voisset C and Andrawiss M (2000). "Retroviruses at a glance." Genome Biol **1**(3).

## W

- Wang-Johanning F, Frost AR, et al. (2003). "Detecting the expression of human endogenous retrovirus E envelope transcripts in human prostate adenocarcinoma." Cancer **98**(1): 187-97.
- Wang-Johanning F, Frost AR, et al. (2001). "Expression of human endogenous retrovirus k envelope transcripts in human breast cancer." Clin Cancer Res **7**(6): 1553-60.
- Wang-Johanning F, Liu J, et al. (2007). "Expression of multiple human endogenous retrovirus surface envelope proteins in ovarian cancer." Int J Cancer **120**(1): 81-90.
- Wang F, L. H., Blanton WP, Belkina A, Lebrasseur NK, Denis GV. (2009). "Brd2 disruption in mice causes severe obesity without Type 2 diabetes." Biochem J. **425**(1): 71-83.
- Wanisch K and Yáñez-Muñoz RJ (2009). "Integration-deficient lentiviral vectors: a slow coming of age." Mol Ther **17**(8): 1316-32.
- Weidhaas JB, Angelichio EL, et al. (2000). "Relationship between retroviral DNA integration and gene expression." J Virol **74**(18): 8382-9.
- Weidner-Glunde M and S. T. Ottinger M (2010 ). "WHAT do viruses BET on?" Front Biosci **15**: 537-49.
- Weiss RA (2006). "The discovery of endogenous retroviruses." Retrovirology **3**(67).
- Witte V, Laffert B, et al. (2004). "HIV-1 Nef mimics an integrin receptor signal that recruits the polycomb group protein Eed to the plasma membrane." Mol Cell **13**(2): 179-90.
- Wu SY and Chiang CM (2007). "The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation." J Biol Chem **282**(18): 13141-5.
- Wu X, Li Y, et al. (2003). "Transcription start regions in the human genome are favored targets for MLV integration." Science. **300**(5626): 1749-51.
- Wu X, Qian K, et al. (2010). "Recombinant avian leukosis viruses of subgroup J isolated from field infected commercial layer chickens with hemangioma and myeloid leukosis possess an insertion in the E element." Vet Res Commun **34**(7): 619-32.

## Y

- Yamashita M and Emerman M (2006). "Retroviral infection of non-dividing cells: old and new perspectives." Virology **344**(1): 88-93.
- Yang F, Leon O, et al. (1999). "Functional interactions of the HHCC domain of moloney murine leukemia virus integrase revealed by nonoverlapping complementation and zinc-dependent dimerization." J Virol. **73**(3): 1809-17.
- Yang F and Roth MJ (2001). "Assembly and catalysis of concerted two-end integration events by Moloney murine leukemia virus integrase." J Virol. **75**(20): 9561-70.
- Yang ZN, Mueser TC, et al. (2000). "Crystal structure of an active two-domain derivative of Rous sarcoma virus integrase." J Mol Biol **296**(2): 535-48.
- Yoo D and Giulivi A (2000). "Xenotransplantation and the potential risk of porcine viruses for xenogeneic transmission." Can J Veterinary Res **64**(4): 193-203.
- Yu P, Zhang P, et al. (2009). "Studies on long-term infection of human cells with Porcine endogenous retrovirus." Acta Virol **53**(3): 169-74.
- Yung E, Sorin M, et al. (2001). "Inhibition of HIV-1 virion production by a transdominant mutant of integrase interactor 1." Nat Med **7**(8): 920-6.

## Z

- Zennou V, Petit C, et al. (2000). "HIV-1 genome nuclear import is mediated by a central DNA flap." Cell. **101**(12): 173-85.
- Zhao Z, McKee CJ, et al. (2008). "Subunit-specific protein footprinting reveals significant structural rearrangements and a role for N-terminal Lys-14 of HIV-1 Integrase during viral DNA binding." J Biol Chem **283**: 5632-5641.
- Zheng R, Ghirlando R, et al. (2000). "Barrier-to-autointegration factor (BAF) bridges DNA in a discrete, higher-order nucleoprotein complex." Proc Natl Acad Sci U S A **97**(16): 8997-9002.
- Zheng R, Jenkins TM, et al. (1996). "Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity." Proc Natl Acad Sci U S A. **93**(24): 13659-64.