



Etude des fonctions cellulaires de SAMHD1, facteur de restriction du VIH-1

Tania Louis

► To cite this version:

Tania Louis. Etude des fonctions cellulaires de SAMHD1, facteur de restriction du VIH-1. Sciences du Vivant [q-bio]. Université de Montpellier, 2015. Français. NNT: . tel-01548610v1

HAL Id: tel-01548610

<https://theses.hal.science/tel-01548610v1>

Submitted on 17 Aug 2015 (v1), last revised 27 Jun 2017 (v2)

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

THÈSE

Pour obtenir le grade de
Docteur

Délivré par **L'Université de Montpellier**

Préparée au sein de l'école doctorale Sciences Chimiques
et Biologiques pour la Santé
Et de l'unité de recherche Institut de Génétique Humaine

Spécialité : **Biologie**

Présentée par **Tania LOUIS**

**Etude des fonctions cellulaires de SAMHD1,
facteur de restriction du VIH-1**

Soutenue le 8 juillet 2015 devant le jury composé de :

M. Moncef BENKIRANE, DR, CNRS
M. Uriel HAZAN, PU, CNRS/ENS Cachan
M. Olivier SCHWARTZ, DR, Institut Pasteur
Mme Clarisse BERLIOZ-TORRENT, DR, INSERM
Mme Nathalie CHAZAL, MCU, CNRS/UM

Directeur de thèse
Rapporteur
Rapporteur
Examinatrice
Examinatrice



Table des matières

Remerciements	4
RÉSUMÉS	6
Résumés de la thèse	6
Résumé français :	6
English Abstract:	7
Résumés de thèse vulgarisés	8
Résumé français :	8
English Abstract:	8
Table des abréviations	9
INTRODUCTION	11
Préambule	12
I/ La virologie, une discipline pionnière	12
A/ Découverte d'activités catalytiques nouvelles : exemple de la transcription inverse	12
1/ Le dogme central de la biologie moléculaire face à la transcription inverse	12
2/ Importance biologique des rétro-éléments	14
B/ Les virus comme moteurs évolutifs	16
1/ Les syncytines ou la création d'une Classe	16
2/ Reproduction sexuée et dynamique du chat du Cheshire	18
C / Découvertes de mécanismes cellulaires fondamentaux : l'export nucléaire	19
1/ Identification de Crm1, une karyophérine majeure	19
2/ Identification de TAP, acteur clé de l'export des ARNm cellulaires	21
D/ Quand les virus deviennent des outils	22
1/ Outils pour les chercheurs : exemple du traçage transneuronal	22
2/ Outils thérapeutiques	23
a/ Production de vecteurs viraux	23
b/ Applications de différents types de vecteurs	25
II/ Compréhension des mécanismes de l'oncogenèse	27
A/ Evolution pathologique des caractéristiques cellulaires	27
1/ Définitions et concepts principaux	27
2/ De la cellule saine à la cellule cancéreuse	29
a/ Le cancer, une accumulation de dysfonctionnements	29
b/ Caractéristiques des cellules cancéreuses et hétérogénéité tumorale	30
B/ Comment identifier les gènes impliqués dans l'oncogenèse ?	32
1/ Oncogènes et suppresseurs de tumeurs	32
2/ Apports de l'étude des cancers familiaux et héréditaires	34
3/ Importance historique des virus dans la compréhension de l'oncogenèse	34
C/ Processus cellulaires perturbés lors de l'oncogenèse	36
1/ Résistance à l'apoptose	37
2/ Perte de contrôle du cycle cellulaire et augmentation de la prolifération	40

3/ Altération de la réparation des dommages à l'ADN	42
III/ SAMHD1, histoire et contexte d'une découverte majeure	47
A/ Le VIH : structure virale et adaptation à l'environnement cellulaire	47
1/ Les VIH, des lentivirus à l'origine d'une pandémie	47
a/ Le SIDA : de la découverte du VIH à la gestion d'une pandémie	47
b/ Organisation des particules et des génomes viraux	47
c/ Classification et origines des VIH	50
2/ Réplication dans les cellules hôtes	51
a/ Cycle réplcatif du VIH	51
b/ Utilisation de facteurs cellulaires facilitants	53
c/ Echappement à l'immunité innée	55
3/ VIH et facteurs de restriction	56
a/ Caractéristiques et fonctions des facteurs de restrictions	56
b/ Mécanismes d'échappement viral	60
B/ SAMHD1, du phénotype à la caractérisation moléculaire	61
1/ Identification d'un nouveau facteur de restriction	61
a/ Importance de Vpx pour l'infection de certains types cellulaires	61
b/ Caractérisation de l'activité anti-VIH de SAMHD1	63
c/ Une activité plus étendue qu'on ne le pensait	64
2/ Caractérisation de deux activités catalytiques	66
a/ SAMHD1 est une dNTP triphosphohydrolase	66
b/ SAMHD1 est une exonucléase	67
c/ Articulation de ces deux activités	68
3/ Une régulation encore peu connue	69
a/ Régulation transcriptionnelle et post-transcriptionnelle	69
b/ Régulation post-traductionnelle	70
4/ Fonction(s) et mécanismes d'action	71
a/ Mécanisme antiviral en cours d'élucidation	71
b/ Fonction(s) cellulaire(s) peu voire non connues	72
Problématique	75
RESULTATS	76
Eléments de contextualisation	77
Article publié	78
Présentation de ma démarche expérimentale	107
DISCUSSION	112
Implications thérapeutiques nouvelles de SAMHD1	113
SAMHD1 et résistances à certaines chimiothérapies	113
Perte de SAMHD1 et tumorigénèse	114
Mise en évidence de fonctions cellulaires de SAMHD1	115
SAMHD1 et progression du cycle cellulaire	115
SAMHD1 et réparation des dommages à l'ADN	117
Fonctions cellulaires et activités catalytiques de SAMHD1	118
Effet de SAMHD1 sur le cycle cellulaire	119

SAMHD1 et réparation des dommages à l'ADN	120
SAMHD1 et réponse immunitaire innée	122
Fonctions de SAMHD1 à l'échelle de l'organisme	123
Méconnaissance des mécanismes de régulation de SAMHD1.....	123
Angles d'approches pour l'étude de SAMHD1	124
SAMHD1, un facteur de restriction	125
SAMHD1, un garant du fonctionnement cellulaire	126
Conclusion générale	126
ANNEXES.....	128
Descriptif des autres projets menés au cours du doctorat.....	129
1/ Etude de la régulation de la protéine Vpx	129
2/ Etude de l'effet de Vpx sur l'expression à long terme de SAMHD1	130
3/ Recherche d'autres ubiquitine ligases recrutées par Vpx	130
Matériel et méthodes supplémentaire	132
Traitements des cellules.....	132
Séquençage	132
Méthodes précédemment décrites.....	132
Echantillons	132
Références bibliographiques	133

Remerciements

Je n'ai jamais été très douée pour les remerciements, et au moment d'écrire ceux-ci je pense surtout à toutes les personnes qui ont été là pour moi pendant ma thèse et que je vais malencontreusement oublier. D'avance, je m'en excuse.

Je tiens tout d'abord à remercier Monsef Benkirane de m'avoir accueillie dans son équipe, pour mon stage de master 2 puis pour ma thèse.

Je remercie également les autres membres de mon jury, Clarisse Berlioz-Torrent, Nathalie Chazal, Uriel Hazan et Olivier Schwartz, qui ont accepté d'évaluer mon travail de doctorat.

J'ai bien sûr une pensée pour tous les membres actuels et passés de l'équipe Benkirane, avec qui j'ai parfois partagé plus qu'un bureau ou une paillasse. Parmi tous, j'adresse un remerciement particulier à Sabine, Estelle et Christelle. Vous avez été à la fois des bouffées d'oxygène et des soutiens solides quand j'en ai eu besoin, j'espère avoir pu vous rendre la pareille. Je compte sur vous pour qu'on garde contact !

Je souhaite remercier l'ensemble des personnels de l'ancienne Université Montpellier 2 qui ont contribué à faire de mon expérience d'enseignement ce qu'elle a été : la découverte d'une passion. Et je remercie également mes étudiants, notamment ceux du groupe A de la série A (deuxième année de licence de biologie, année 2014-2015). Vous ne lirez probablement jamais ce message, mais l'ambiance de travail avec vous était vraiment extra-ordinaire.

J'adresse un remerciement tout particulier au personnel de Genopolys, notamment à Magali et Géraldine qui, sans le savoir, ont vraisemblablement changé ma vie. Merci de m'avoir donné, à l'occasion d'une recherche de bénévoles pour la Fête de la Science, l'occasion de découvrir ce que pouvait être la médiation scientifique. Et surtout merci de m'avoir soutenue et de m'avoir fait confiance : sans vous, Galilée 2.0 n'aurait sans doute jamais vu le jour, et en tous cas n'aurait jamais pu grandir comme il l'a fait.

J'ai évidemment envie de remercier toutes les personnes qui, au fil des années, m'ont aidée à conserver une vie personnelle équilibrée au-delà de la thèse et qui ont parfois su prêter une oreille attentive et constructive à mes errances.

Merci aux N'Impro (et assimilés) dont plusieurs centaines de kilomètres n'ont pas réussi à me séparer. Merci à la SSC qui même éparpillée aux quatre coins du monde résiste encore et toujours (et commence même à se multiplier, bonjour Inaya !). Merci à l'incomparable promo des « jeudis Pasteur », j'espère qu'on arrivera à se retrouver un peu une fois que la vague des soutenances sera retombée. Merci aux membres actuels et passés de l'asso CBS². Merci aux promotions de TAUSTiens qui se reconnaîtront. Merci à tous ceux que je ne peux pas ranger dans des cases mais qui ont quand même tellement compté (c'est à ce moment précis que je suis certaine d'oublier des gens importants... permettez-moi d'esquiver lâchement la question en évitant de faire une liste qui ne serait pas exhaustive).

Et bien qu'ils soient arrivés relativement tard par rapport à la durée totale de ma thèse, je veux adresser un remerciement très spécial à l'ensemble des comédiens qui ont été assez fous pour me suivre dans le projet Galilée 2.0. Merci. Que vous ayez simplement participé à la pièce, que vous m'ayez activement aidée dans l'écriture de ce manuscrit (j'ai vraiment une pensée particulière pour toi, Stan), que nous ayons partagé de longues discussions sur le fonctionnement actuel de la recherche, vous m'avez tous beaucoup apporté. Je pense que, dans ma tête, je resterai longtemps votre « miseuse en scène ». Et j'en suis particulièrement fière.

Enfin, et ça ne sera pas très original, je tiens à remercier mes parents, pour leur soutien permanent dans ce périple étrange qu'ont été mes études supérieures.

Je me suis forcée à garder le meilleur pour la fin (en réalité j'ai craqué et ce paragraphe a été rédigé juste après le premier de cette partie, mais chut c'est un secret). Je ne sais pas si tu réalises quelle force peut avoir ce mot, mais, Julian : merci. Merci d'être à mes côtés depuis plus de huit ans. Merci de m'avoir supportée dans les moments difficiles (je sais à quel point je peux être pénible quand je suis stressée et je sais à quel point mes efforts pour être plus vivable sont peu efficaces). Merci de m'avoir écoutée dans les périodes de grand doute. Merci d'avoir égayé mon quotidien, simplement en étant toi-même. Merci de m'avoir poussée à continuer d'avancer, même quand le chemin tout tracé a laissé place aux broussailles. Merci d'avoir vérifié mes « to do-lists » pour lutter contre ma tendance naturelle à la procrastination. Merci de m'avoir offert un violon, qui n'a pas vraiment aidé à lutter contre ma tendance naturelle à la procrastination. Merci pour beaucoup de choses. Et entre toutes, merci de savoir qui je suis et de m'aimer quand même.

Résumés

Résumés de la thèse

Résumé français :

L'étude des interactions entre un pathogène et son hôte, bien qu'ayant généralement pour objectif de contrôler l'infection par le pathogène, permet parfois de découvrir des éléments fondamentaux sur le fonctionnement de l'hôte. J'ai choisi d'étudier les fonctions cellulaires d'une protéine initialement identifiée comme un facteur de restriction du VIH-1.

SAMHD1 (SAM domain and HD domain-containing protein 1) est une protéine exprimée dans la plupart des tissus humains. Elle est capable d'hydrolyser les déoxyribonucléotides triphosphates (dNTP) cellulaires et possède une activité nucléase ciblant différents acides nucléiques dont les ARN simple brin *in vitro*.

Des mutations dans le gène SAMHD1 entraînent le développement d'une maladie auto-immune pouvant conduire à la mort précoce des nourrissons, ce qui suggère un rôle de la protéine correspondante dans la régulation de la réponse immunitaire. Il a été montré que SAMHD1 est un facteur de restriction capable d'empêcher l'infection de cellules ne se divisant pas par le VIH-1. La protéine virale Vpx, exprimée par le VIH-2, est capable d'induire la dégradation de SAMHD1 par le protéasome et permet de rendre permissives les cellules initialement résistantes à l'infection par le VIH.

SAMHD1 est en réalité capable de restreindre l'infection par des virus aussi différents que les rétrovirus et le virus de l'herpès simplex 1. Néanmoins, le mécanisme permettant à SAMHD1 de contrecarrer différents virus reste aujourd'hui sujet à controverse. Initialement considéré comme agissant en dégradant les dNTP cellulaires, SAMHD1 semble également capable de dégrader l'ARN génomique du VIH-1.

Si de nombreux travaux portent sur l'activité antivirale de SAMHD1, peu de données sont disponibles concernant la fonction cellulaire de cette protéine. Or SAMHD1 est capable de réguler la quantité de dNTP cellulaires et d'interagir avec certains acides nucléiques. Ces données font de SAMHD1 un acteur potentiel de différents processus cellulaires fondamentaux sensibles à la quantité intracellulaire de dNTP, notamment la réplication du génome ou la réparation des dommages à l'ADN.

J'ai montré au cours de mon doctorat que SAMHD1 module le cycle cellulaire et notamment que la surexpression de cette protéine ralentit la prolifération cellulaire. J'ai également observé que la surexpression de SAMHD1 augmente la sensibilité des cellules aux agents induisant des ruptures double brin de l'ADN. De plus, j'ai découvert qu'en cas de ruptures double brin de l'ADN cellulaire, SAMHD1 est régulé de façon spécifique par phosphorylation sur sa thréonine 592 et est recruté aux sites de cassures.

D'autres travaux ont confirmé l'importance de la régulation de SAMHD1 au cours du cycle cellulaire, sa surexpression et sa réduction induisant toutes deux un ralentissement de la prolifération cellulaire. En complément de mes résultats, quelques études suggèrent que SAMHD1 joue un rôle dans le maintien de l'intégrité du génome, qui pourrait être dû à son effet sur la réponse aux dommages à l'ADN. Dans l'ensemble, ces résultats font de SAMHD1 un garant de l'homéostasie cellulaire.

J'ai de plus montré que l'expression de SAMHD1 est réduite chez environ 80% des patients souffrant de leucémie lymphoïde chronique. La perte de cette protéine est donc corrélée à l'apparition d'une maladie découlant de la perturbation du fonctionnement cellulaire. L'étude d'échantillons d'autres types de tumeurs montre que, dans de moindres proportions, l'altération de l'expression de SAMHD1 est une caractéristique générale des cancers.

Mes travaux de doctorat soulignent ainsi le rôle fondamental de SAMHD1 dans le maintien de l'intégrité cellulaire.

English Abstract:

Understanding host pathogen interactions reveals not only important information regarding the replication cycle of the pathogen but it often leads to the discovery and better understanding of key biological processes of the host. The aim of my PhD was to decipher the cellular functions of the HIV-1 restriction factor SAMHD1.

SAMHD1 (SAM domain and HD domain-containing protein 1) is expressed in most human tissues. This protein is able to hydrolyze cellular deoxyribonucleotides triphosphate (dNTP) and possesses a nuclease activity primarily against single stranded RNA.

Mutations in SAMHD1 have been described in patients suffering from an auto-immune disease causing premature death of newborns. This phenotype suggests a role of SAMHD1 in the control of immune response. Moreover, SAMHD1 restricts HIV-1 in non-cycling cells. The HIV-2 accessory protein Vpx induces SAMHD1 degradation by the proteasome, conferring cell permissiveness to HIV.

In fact, the antiviral activity of SAMHD1 has been extended to other viruses including Herpes Simplex Virus 1 and Hepatitis B virus. Nevertheless, the mechanism by which SAMHD1 restrict HIV replication is debated. It was initially thought to act by depleting the dNTP pool but recent studies highlighted a potential role of SAMHD1 nuclease function in degrading HIV-1 genomic RNA.

Many studies aiming at understanding the antiviral activity of SAMHD1 are being pursued, whereas little is known about the cellular function of this protein. The fact that SAMHD1 is able to regulate the cellular dNTP pool and to interact with nucleic acids suggests a key role of this protein in cellular processes, such as DNA replication and repair.

During my PhD, I showed that SAMHD1 modulates the cell cycle, as the overexpression of this protein slows down cell proliferation. I also observed that SAMHD1 overexpression increases cellular sensitivity to double strand DNA breaks-inducing agents. Moreover I discovered that, after double strand breaks induction, SAMHD1 is specifically regulated by phosphorylation on its threonine 592 and recruited at the damaged sites.

Other studies confirmed the importance of SAMHD1 regulation along the cell cycle as its overexpression and depletion both decrease cell proliferation. In addition to my observations, some studies suggested that SAMHD1 is important to maintain genomic integrity, presumably through its implication in DNA repair. Altogether, these results promote SAMHD1 as a key player in cellular homeostasis.

I additionally showed that SAMHD1 expression is reduced in 80% of patients suffering from chronic lymphocytic leukemia (CLL). SAMHD1 loss is therefore correlated to the development of a disease due to disturbances of cellular integrity. Looking at samples from different types of tumors, I showed that SAMHD1 loss is shared between all tested cancers, although at lesser extent than in CLL.

My PhD work underlines the central role of SAMHD1 to maintain cellular integrity.

Résumés de thèse vulgarisés

Résumé français :

Les virus utilisent nos cellules pour se multiplier. Ainsi, en cherchant à comprendre comment les bloquer, on fait parfois des découvertes fondamentales sur le fonctionnement des cellules humaines.

J'ai étudié SAMHD1, une protéine capable d'empêcher le virus du SIDA d'infecter certaines de nos cellules, pour connaître ses autres fonctions. J'ai d'abord découvert qu'une augmentation de la quantité de SAMHD1 ralentit la vitesse à laquelle nos propres cellules se multiplient et favorise leur mort si leur matériel génétique est trop abimé.

Or les cellules cancéreuses sont connues pour leurs capacités anormales de multiplication et de survie. J'ai donc mesuré la quantité de SAMHD1 dans des cellules de tumeurs et j'ai observé que SAMHD1 avait disparu de certaines d'entre elles.

Ces résultats montrent que SAMHD1 n'est pas seulement capable de bloquer le virus du SIDA mais est nécessaire au bon fonctionnement de nos cellules et pourrait éventuellement bloquer le développement des cancers.

English Abstract:

Viruses are able to use our cells to multiply themselves. This is why we sometimes discover interesting things about our own biology while trying to find a way to block viruses.

I have studied SAMHD1, a protein able to protect some of our cells from the AIDS virus, to uncover its others functions. First, I discovered that an increase of SAMHD1 quantity slows down the multiplication rate of our own cells and promotes their death if their genetic material is too damaged.

As cancer cells are known for their abnormal abilities to multiply and survive, I measured SAMHD1 amounts in cancerous cells from patients. I then observed that SAMHD1 had disappeared from some of them.

These results show that SAMHD1 is not only able to block the AIDS virus but is also required for the proper functioning of our cells, and may be able to prevent cancers development.

Table des abréviations

AAV : virus associés aux adénovirus	IL : interleukine
ADN : acide désoxyribonucléique	IN : intégrase
ADNc : ADN complémentaire	IRES : site d'entrée interne des ribosomes
AGS : syndrome Aicardi-Goutière	IRF : interferon regulatory factor
ARN : acide ribonucléique	ISG : gènes stimulés par l'interféron
ARNm : ARN messenger	ITAF : facteurs trans-activateurs des IRES
ARNt : ARN de transfert	ITR : répétition terminale inversée
ARNr : ARN ribosomal	LINE : long élément nucléaire intercalé
CaMV : virus de la mosaïque du chou-fleur	LLA-B : leucémie lymphoïde aigue B
CDC : center for disease control	LLA-T : leucémie lymphoïde aigue T
CDK : kinase dépendante des cyclines	LLC : leucémie lymphoïde chronique
cnLOH : perte d'hétérozygotie neutre en nombre de copies	LMA : leucémie myéloïde chronique
CPT : camptothécine	LMC : leucémie myéloïde aigue
CRL4 : culline ring ligase 4	lncARN : long ARN non codant
CTE : constitutive transport element	LTR : longue répétition terminale
dATP : désoxyadenosine triphosphate	MA : matrice
dCTP : désoxycytidine triphosphate	MAPK : mitogen-activated protein kinase
DD : domaines de mort	miR : microARN
dGTP : désoxyguanosine triphosphate	MLV : virus de la leucémie murine
dNTP : désoxynucléoside triphosphate	MMC : mitomycine C
dTTP : désoxythymidine triphosphate	MMS : méthanesulfonate de méthyle
EIAV : virus de l'anémie infectieuse des équidés	MPMV : virus Mason-Pfizer du singe
GDP : guanosine diphosphate	NC : nucléocapside
GPI : glycosylphosphatidylinositol	NHEJ : jonction d'extrémités non homologues
GTP : guanosine triphosphate	NTP : nucléoside triphosphate
GTPase : GTP hydrolase	ORF : cadre ouvert de lecture
HERV : retrovirus endogène humain	PAMP : motif moléculaire associé aux pathogènes
HSV-1 : virus herpes simplex 1	PBMC : cellules mononuclées du sang périphérique
HTLV-1 : virus T-lymotropique humain 1	PIC : complexe de pré-intégration
IAP : protéines inhibant l'apoptose	PR : protéase
ires : interféron	PRR : récepteur reconnaissant les pathogènes

PRRSV : virus du syndrome reproducteur et respiratoire porcin

pTEFb : positive transcription elongation factor b

qPCR : réaction de polymérisation en chaîne quantitative

RNR : ribonucléotide réductase

RRE : élément de réponse à Rev

RT : transcriptase inverse

SAM : sterile alpha motif

SCF : skp, culline, F-box

SEC : complexe de super élongation

shARN : petit ARN en épingle à cheveux

SIDA : syndrome d'immunodéficience acquise

SINE : petit élément nucléaire intercalé

siRNA : ARN interférent

snARN : petit ARN nucléaire

SV40 : virus simien 40

T592 : thréonine 592

TAR : élément de réponse à la transactivation

TLR : récepteur Toll-like

uORF : ORF en amont

UTR : région non traduite

VHB : virus de l'hépatite B

VHC : virus de l'hépatite C

VIB : virus de l'immunodéficience bovin

VIF : virus de l'immunodéficience félin

VIH : virus de l'immunodéficience humain

VIS : virus de l'immunodéficience simien

VLA : virus de la leucémie aviaire

VMT : virus de la mosaïque du tabac

VSV : virus de la stomatite vésiculaire

Introduction

Préambule

A chaque règle énoncée en biologie correspond au moins un virus capable de la contredire. Cette observation va des règles générales telles que « l'acide désoxyribonucléique (ADN) est le support de l'information génétique (Alberts et al., 2008) » aux règles plus spécifiques, y compris celles concernant les virus eux-mêmes et notamment leur taille maximale qui ne cesse d'être augmentée par la découverte de virus de plus en plus volumineux (Etten et al., 2010). Les virus sont de tels briseurs de dogmes biologiques qu'ils suscitent de nombreuses discussions sur la définition même du « vivant » (Forterre, 2012).

Cet aspect atypique et cette diversité sans limites apparentes sont parmi les intérêts majeurs de la virologie sur les plans intellectuel et scientifique. A cela s'ajoute le fait que les virus sont dépendants de cellules hôtes infectées pour se multiplier. Ainsi, l'étude de la réplication d'un virus est souvent intimement liée à l'étude de la cellule qu'il infecte et la recherche sur les interactions hôtes-pathogènes est à l'origine de découvertes majeures dans le champ de la biologie cellulaire.

Tout cela fait de la virologie une discipline pionnière en biologie, à l'origine de nombreuses découvertes fondamentales, de nombreux concepts novateurs et de plus en plus d'outils moléculaires.

Mon doctorat a été consacré à l'étude de SAMHD1 (SAM domain and HD domain-containing protein 1), un facteur de restriction du Virus de l'Immunodéficience Humaine (VIH) identifié en 2011. En effet, peu d'informations étaient alors disponibles sur cette protéine au-delà de son activité antivirale. Mes travaux se sont orientés vers la caractérisation des fonctions cellulaires de SAMHD1.

Cela m'a conduit à m'intéresser à différents processus cellulaires susceptibles de réguler ou d'être régulés par SAMHD1, parmi lesquels on trouve essentiellement des processus liés aux acides nucléiques tels que la réplication et la réparation des dommages à l'ADN.

J'ai également travaillé sur trois autres projets indépendants qui ne seront pas présentés dans ce manuscrit. Ils concernent la régulation de SAMHD1, par la protéine Vpx ou par d'autres facteurs, et la régulation de la protéine Vpx elle-même.

I/ La virologie, une discipline pionnière

A/ Découverte d'activités catalytiques nouvelles : exemple de la transcription inverse

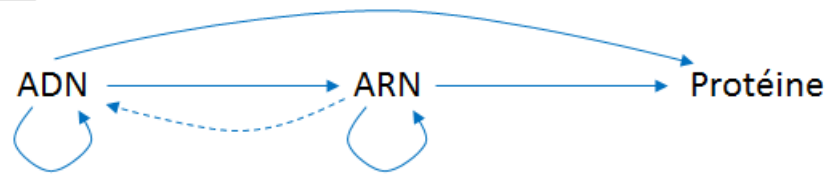
1/ Le dogme central de la biologie moléculaire face à la transcription inverse

Le dogme central de la biologie moléculaire a été énoncé par Francis Crick dans une note datée de 1956 (Crick, 1956), en tant que concept non étayé par des preuves expérimentales solides. Il y fait à nouveau référence dans une publication datée de 1958 (Crick, 1958), rassemblant les données expérimentales disponibles à l'époque concernant la synthèse des protéines. Ce dogme est généralement énoncé de la façon suivante : « l'information génétique, ou gène, sous forme d'ADN peut être transcrite sous forme d'une molécule d'acide ribonucléique (ARN) mobile qui sera à son tour traduite pour donner une protéine ». En réalité, Crick avait une vision moins affirmative et s'inscrivait dans une logique de réflexion sur le transfert d'information dans les cellules. Il insistait essentiellement sur le fait que d'après lui l'information pouvait passer des acides nucléiques aux protéines mais ne pouvait pas effectuer le chemin inverse.

Dans son premier manuscrit (Crick, 1956), il considère l'ARN uniquement en tant que support d'information pour la traduction des protéines et précise que les ARN néosynthétisés sont rapidement dégradés s'ils ne sont pas associés à la machinerie de traduction (encore inconnue à l'époque). Dès

cette publication, il évoque la possibilité de synthèse d'ARN à partir d'ARN et suggère sans détailler son propos la possibilité d'un transfert d'information de l'ARN à l'ADN, bien que celui-ci lui paraisse moins probable que le transfert en sens inverse (Figure 1).

On pourrait observer :



Mais jamais :

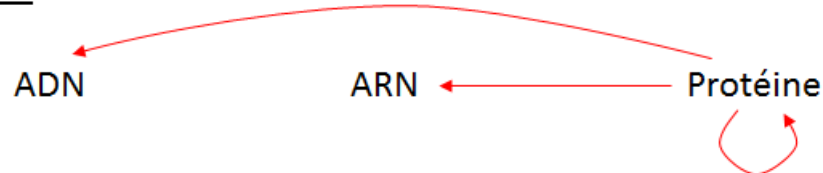


Figure 1 : schéma d'après F. Crick, en 1956 (Crick, 1956). Dans cette première formulation du dogme central de la biologie moléculaire, le transfert d'information est envisagé comme possible (flèches pleines bleues), possible mais peu probable (flèche bleue pointillée) ou impossible (flèches rouges pleines) entre certains types de biomolécules.

Dès son article de 1958 (Crick, 1958), Crick utilise des expériences réalisées sur des virus pour soutenir son raisonnement sur la nature de l'ARN. Il fait notamment référence aux travaux de Fraenkel-Conrat. Ce dernier a montré que des plantes inoculées simultanément avec l'ARN d'une première souche de virus de la mosaïque du tabac (VMT) et les protéines d'une seconde souche de VMT produisent un virus identique à celui de la première souche (Fraenkel-Conrat, 1956). Ce sont donc déjà des études de virologie qui ont amené Crick à considérer l'ARN comme un support d'information dirigeant la synthèse protéique.

Mais la virologie est également à l'origine du premier grand chamboulement de ce dogme puisqu'elle a permis de montrer que le transfert d'information de l'ARN vers l'ADN, que Crick estimait initialement peu probable, était bel et bien possible. En effet, en 1970, Baltimore a montré que les particules virales de deux virus à ARN oncogènes, le virus du sarcome de Rous et le virus de la leucémie murine de Rauscher, contiennent une enzyme capable de synthétiser de l'ADN à partir d'ARN : la transcriptase inverse (Baltimore, 1970). Dans la même édition de la revue *Nature*, Temin et Mizutani, qui postulaient l'existence d'une telle activité catalytique depuis plusieurs années, ont publié des résultats similaires concernant le virus du sarcome de Rous et le virus de la myéloblastose aviaire (Temin and Mizutani, 1970).

Ainsi, lorsque Crick reformule le dogme central de la biologie moléculaire en 1970 (Crick, 1970), il considère comme des mécanismes généraux les transferts d'information de l'ADN vers l'ADN, de l'ADN vers l'ARN et de l'ARN vers les protéines et décrit une nouvelle classe de « transferts spéciaux », n'ayant pas lieu dans la plupart des cellules mais restant possibles dans des conditions particulières. Il y inclut notamment les transferts d'information de l'ARN vers l'ARN (Baltimore and Franklin, 1962; Baltimore et al., 1963) et de l'ARN vers l'ADN tout en rappelant que ceux-ci n'ont été observés que dans des cellules infectées par certains virus.

En moins de quinze ans, la découverte par des virologues d'activités enzymatiques de type transcriptase inverse, capables de synthétiser de l'ADN à partir d'ARN, est venue nourrir la formulation

d'une théorie de référence en biologie moléculaire. Cette découverte fondamentale a d'ailleurs valu le Prix Nobel de médecine à ses auteurs David Baltimore et Howard Martin Temin en 1975, prix partagé avec Renato Dulbecco.

2/ Importance biologique des rétro-éléments

Assez rapidement après la découverte des transcriptases inverses virales, des traces d'activités comparables ont été détectées dans les génomes de Mammifères sous la forme de rétrotransposons (Hollis et al., 1982; Nishioka et al., 1980; Sharp, 1983) et de rétrovirus endogènes (Lueders and Kuff, 1977). Les rétro-éléments sont définis comme des éléments mobiles du génome dont la duplication nécessite un intermédiaire sous forme d'ARN servant de support à une transcriptase inverse pour permettre la synthèse d'une copie ADN qui sera intégrée dans un nouveau locus (Figure 2, partie de gauche).

Parmi ceux-ci se trouvent les rétrovirus endogènes. Ces éléments ont une structure identique à celle du génome des rétrovirus mais ont généralement accumulé des mutations inactivant certains de leurs gènes ce qui les empêche de se multiplier comme les rétrovirus dans le cadre d'une infection. Certains rétrovirus endogènes, comme l'élément HERV-K (Human Endogenous RetroVirus-K) chez l'Homme, sont néanmoins capables de conduire à la production de particules virales mais ces dernières ne sont pas infectieuses (Bieda et al., 2001).

Le second type de rétro-éléments, les rétrotransposons, est subdivisé en deux grandes familles. Les rétrotransposons à LTR (Long Terminal Repeats) ont une structure proche de celle du génome des rétrovirus mais ne possèdent pas de gène Env et ne sont donc pas capables de permettre la production de particules virales. Les rétrotransposons sans LTR ont une structure plus simple. Certains, comme les LINEs, codent leur propre transcriptase inverse alors que d'autres, comme les SINEs, sont dépendants de la transcriptase inverse d'autres rétro-éléments (Tableau 1).

Intermédiaire ADN		Transposons	ITR Transposase ITR		3%	
Intermédiaire ARN	- LTR	Rétrotransposons sans LTR	LINEs	P ORF1 ORF2 AAAAAAA	21%	
			SINEs	P SNC AAAAAAA	13%	
	+ LTR	Rétrotransposons à LTR	LTR Gag Pol LTR			8%
		Rétrovirus endogènes	LTR Gag Pol Env LTR			

Tableau 1 : récapitulatif des principales classes d'éléments mobiles retrouvés dans le génome humain. La dernière colonne indique la proportion du génome humain constituée par le type d'élément considéré.

ITR : Répétitions Terminales Inversées, P : promoteur, SNC : Séquence Non Codante,
 LTR : Longues Répétitions Terminales, ORF : cadre ouvert de lecture,
 Gag/Pol/Env/transposase : gènes correspondant aux ORF identifiés.

Le génome humain comprend également des éléments transposables (Liebermann et al., 1986) dont la duplication ne nécessite pas le passage par un intermédiaire ARN : les transposons (Figure 2, partie de droite). Identifiés chez le maïs par Mc Clintock dès 1951 (McClintock, 1951), ces

éléments sont directement excisés du génome et insérés à un nouveau locus par une enzyme de type transposase (Tableau 1).

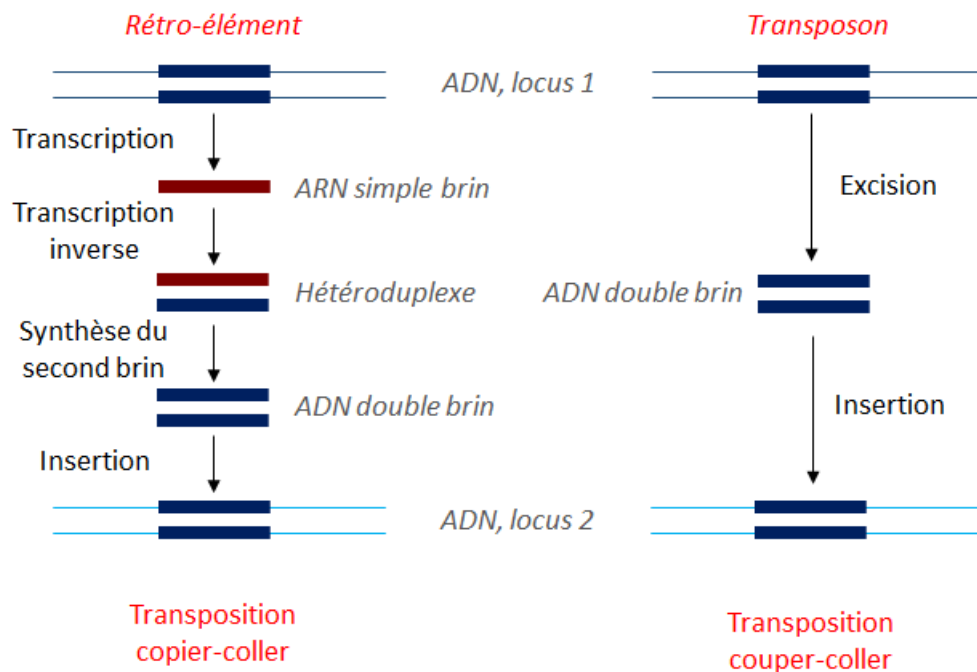


Figure 2 : schéma récapitulant les acides nucléiques intermédiaires impliqués dans le mécanisme de transposition des rétro-éléments et des transposons.

A l'heure actuelle, on considère que le génome humain contient 3% de transposons, 8% de rétrovirus endogènes et de rétrotransposons à LTR et plus de 34% de rétrotransposons sans LTR, ce qui revient à dire que l'ensemble des éléments mobiles représente environ 45% du génome (Consortium, 2002). Si moins de 2% du génome humain contiennent des séquences codant des protéines, il est désormais bien connu que les 98% restants ne sont pas de l'ADN « poubelle » et possèdent un certain nombre de fonctions (Palazzo and Gregory, 2014). Ainsi, l'impact des éléments mobiles sur les génomes n'est pas seulement quantitatif, il est également fonctionnel. En effet, l'insertion d'éléments mobiles dans les gènes cellulaires est susceptible de modifier leur expression selon différents mécanismes (Alberts et al., 2008). En apportant un codon stop au milieu d'un gène, les éléments transposables peuvent empêcher l'expression de la protéine correspondante. Ils peuvent également perturber l'épissage des ARN messagers et conduire à des rétentions d'introns ou des pertes d'exons qui aboutissent à la production d'une protéine nouvelle (Cordaux and Batzer, 2009). Ainsi, en agissant directement sur la séquence codante, les éléments mobiles du génome peuvent conduire à des pertes ou des gains de fonction des gènes cellulaires.

Mais l'insertion d'éléments transposables peut également modifier l'expression des gènes cellulaires sans altérer la séquence codante, via une régulation quantitative de la transcription. En effet, certains rétro-éléments, notamment ceux contenant des LTR, ont une action promotrice sur l'expression des gènes (Cohen et al., 2009). En s'insérant en amont d'une séquence codante, ils vont donc augmenter sa transcription, que celle-ci corresponde à un gène entier ou à une portion de gène.

Les éléments mobiles, constitués à plus de 90% de rétro-éléments chez l'Homme, sont donc des architectes clés des génomes, y compris des génomes de Mammifères. Susceptibles d'altérer le fonctionnement de la cellule, ils sont finement régulés par celle-ci en permanence, au niveau transcriptionnel et au niveau post-transcriptionnel (Rowe and Trono, 2011).

B/ Les virus comme moteurs évolutifs

1/ Les syncytines ou la création d'une Classe

L'intégration d'éléments viraux dans un génome cellulaire est susceptible de modifier son expression et de causer des pathologies mais elle constitue également un moteur évolutif important via l'apparition de nouvelles protéines possédant de nouvelles fonctions potentiellement profitables pour l'hôte.

L'exemple le plus marquant de fonction apportée par l'intégration de séquences virales est vraisemblablement celui ayant conduit à l'apparition de la placentation au cours de l'évolution des Tétrapodes. En 2000, des chercheurs ont observé que la protéine d'enveloppe du rétrovirus endogène défectif HERV-W est très fortement exprimée dans le placenta, notamment au niveau du syncytiotrophoblaste, une structure syncytiale ayant l'épaisseur d'une cellule unique qui forme la couche externe du placenta humain. Ils ont montré que cette protéine permet la fusion de la membrane plasmique des cellules qui l'expriment avec d'autres cellules ou avec des liposomes (Mi et al., 2000). Les propriétés fusogènes de cette protéine et son profil d'expression très particulier suggèrent qu'elle pourrait jouer un rôle dans la fusion des cytotrophoblastes pour former le syncytiotrophoblaste, c'est pourquoi elle a été appelée syncytine (Mi et al., 2000).

Il a par la suite été montré que l'expression de cette protéine d'origine rétrovirale était contrôlée par un facteur de transcription spécifique du placenta : GCMa (Yu et al., 2002). Cela s'explique par le fait que le site d'intégration du rétrovirus ancestral est localisé en aval de deux sites de fixation de GCMa. De plus, une deuxième syncytine a par la suite été décrite chez l'Homme. Comme la première, elle est spécifiquement exprimée dans le placenta, possède des capacités fusogènes, et correspond au gène d'enveloppe encore fonctionnel d'un rétrovirus endogène, HERV-FRD (Blaise et al., 2003).

Sur le plan évolutif, il est intéressant de noter que, contrairement à la plupart des gènes d'enveloppe des rétrovirus endogènes, l'étude des séquences des gènes des syncytines 1 et 2 ne révèle pas de polymorphisme. Cette conservation de séquence est le signe d'une sélection positive forte, suggérant que les protéines correspondantes ont une fonction biologique essentielle. Cependant, ces deux gènes n'ont été retrouvés que chez une partie des Mammifères : les Primates Simiens (de Parseval et al., 2005). Néanmoins, d'autres données expérimentales ont conduit à supposer que des événements d'insertion différents mais du même type sont également à l'origine de l'apparition du placenta chez les autres Mammifères.

Il s'avère en effet que des gènes comparables aux syncytines 1 et 2 des Primates ont été trouvés chez d'autres espèces. Ainsi, chez la souris, deux gènes d'enveloppe présents sous formes de copies uniques et non apparentés aux rétrovirus endogènes murins connus codent des protéines exprimées dans le syncytiotrophoblaste et possédant des propriétés fusogènes : les syncytines A et B (Dupressoir et al., 2005). Des orthologues de ces deux gènes ont été identifiés chez l'ensemble des Muridés mais pas chez les autres Rongeurs. Ils présentent des marqueurs de sélection positive suggérant leur importance fonctionnelle. D'autres travaux menés sur des modèles murins ont permis de réellement mettre en évidence le rôle des protéines correspondantes dans la formation du placenta chez la souris (Dupressoir et al., 2011, 2009).

En plus des Primates et des Rongeurs, des syncytines ont été identifiées chez les Léporidés (non partagée par les autres Lagomorphes) (Heidmann et al., 2009), les Carnivores (Cornelis et al., 2012), les Ruminants (Cornelis et al., 2013), les Tenrecidés (Cornelis et al., 2014), une famille du super-ordre des Afrothériens, et certains Marsupiaux (Cornelis et al., 2015) (Figure 3).

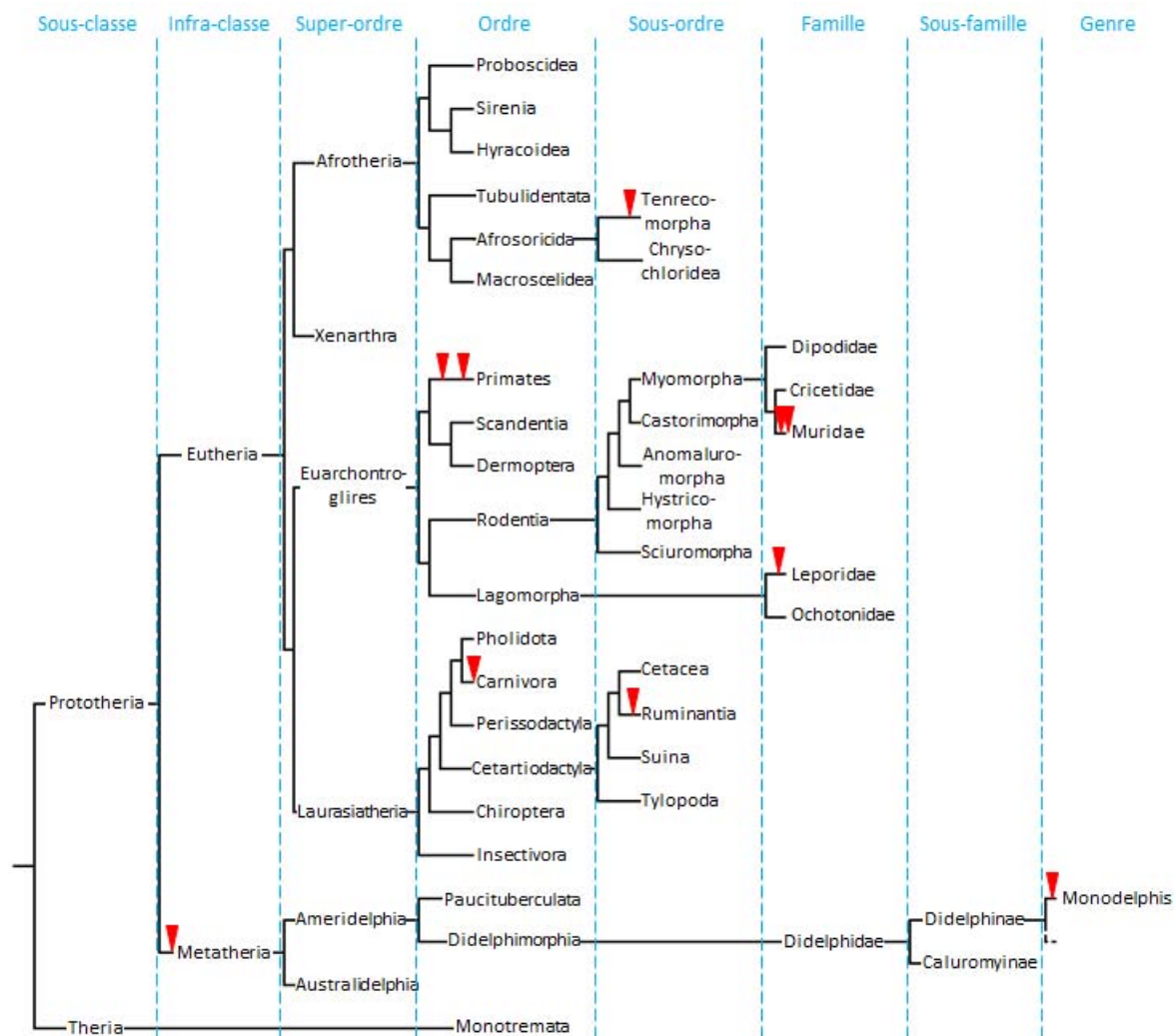


Figure 3 : arbre phylogénétique des Mammifères récapitulant les événements d'insertion de syncytines actuellement répertoriés. Ces derniers sont symbolisés par des triangles rouges. Pour ne pas surcharger la figure, seuls les clades au sein desquels ces insertions ont eu lieu sont détaillés dans la structure de l'arbre (Cornelis et al., 2015; Romanenko et al., 2012; Springer and Murphy, 2007).

L'ensemble de ces découvertes montre que la récupération de gènes d'enveloppe viraux a joué un rôle fondamental dans l'apparition du placenta et donc de la classe des Mammifères. Néanmoins un certain nombre de questions restent encore sans réponse.

Tout d'abord, il paraît surprenant que la protéine à l'origine de la capacité fusogène des cellules placentaires puisse avoir des origines indépendantes chez différents Mammifères. Cela suggérerait que l'apparition du placenta puisse être une forme de convergence évolutive, ce qui n'est pas cohérent avec l'ensemble des données phylogénétiques concernant l'évolution des Mammifères. L'hypothèse la plus probable est qu'une syncytine « originelle » ait été récupérée par l'ancêtre commun des Mammifères mais que les rétentions indépendantes d'autres syncytines au cours de l'évolution aient conduit à sa disparition progressive, si bien qu'elle n'est plus facilement identifiable aujourd'hui (Dupressoir et al., 2012). La découverte de trois gènes de type syncytine à des degrés d'exaptation différents et témoignant de trois insertions différentes au cours du temps chez les Marsupiaux va dans le sens de cette hypothèse (Cornelis et al., 2015).

Dans cette logique d'évolution divergente des mécanismes de formation placentaire après l'apparition du placenta, des travaux sont actuellement menés pour essayer de déterminer dans quelle mesure les différences physiologiques entre les placentas des différents Mammifères peuvent être

expliquées par l'implication de différentes syncytines dans leur formation. Par exemple, chez l'Homme, les syncytines 1 et 2 sont co-exprimées dans une enveloppe syncytiotrophoblastique unique alors que chez la souris les syncytines A et B sont exprimées séparément, dans deux enveloppes syncytiotrophoblastiques différentes (Dupressoir et al., 2011, 2012).

Enfin, la fonction originelle ayant conduit à la sélection des syncytines après leur insertion dans le génome viral est elle aussi sujette à discussion. En effet, il est connu que les protéines d'enveloppe des rétrovirus peuvent également posséder une activité immunosuppressive (Schlecht-Louf et al., 2010) et il s'avère que le domaine responsable de cette activité est conservé chez l'ensemble des syncytines identifiées à ce jour. Cela soulève la question suivante : est-ce que les syncytines ont été sélectionnées pour leur activité fusogène ou pour leur activité immunosuppressive (Dupressoir et al., 2012) ? En effet, une des particularités de la gestation est que l'organisme de la mère doit pouvoir tolérer l'organisme du fœtus, qui devrait être reconnu comme un corps étranger par le système immunitaire maternel. On peut ainsi considérer que du point de vue évolutif, l'apparition d'un mécanisme d'immunosuppression locale est essentiel à l'apparition de la gestation in utero, avant même la mise en place d'un organe comme le placenta. Le fait que la syncytine la plus ancienne observée chez les Marsupiaux, qui connaissent une période de gestation in utero en l'absence de placenta développé, ne soit pas fusogène mais possède un domaine immunosuppresseur fonctionnel et conservé va dans le sens de cette hypothèse (Cornelis et al., 2015). Les protéines d'origines virales, sélectionnées pour leurs fonctions immunosuppressives, auraient alors déterminé la structure du placenta des Mammifères par leurs capacités fusogènes.

2/ Reproduction sexuée et dynamique du chat du Cheshire

Des travaux menés sur des organismes unicellulaires ont amené des chercheurs à supposer que le rôle des virus dans la reproduction dépasse largement la classe des Mammifères puisqu'ils pourraient être à l'origine de la reproduction sexuée.

Cette hypothèse vient de l'étude d'un coccolithophoridé, c'est-à-dire d'une algue unicellulaire pélagique entourée par des plaques de calcite portant le nom de coccolithes : *Emiliana huxleyi*. Les coccolithophoridés jouent un rôle essentiel dans le cycle du carbone. En effet, ils piègent le CO₂ atmosphérique dans leurs coccolithes qui, notamment lors de la mort de l'algue, finissent par se déposer sur les fonds marins où ils sédimentent pour donner de la craie. Le rôle potentiel de ces organismes dans la réponse environnementale à l'augmentation actuelle des gaz à effet de serre est d'ailleurs fortement débattu (Iglesias-Rodriguez et al., 2008; Riebesell et al., 2000).

Ces coccolithophoridés sont à l'origine d'efflorescences algales visibles depuis l'espace car leur protection de calcite est capable de réfracter la lumière et qu'elles peuvent s'étendre sur plusieurs dizaines de milliers de kilomètres carrés. *Emiliana huxleyi* est l'espèce de coccolithophoridé la plus abondante numériquement, on la retrouve dans une grande partie des eaux du globe. Une meilleure connaissance de ce micro-organisme particulier pourrait ainsi permettre d'améliorer les modèles de perturbation du cycle du carbone.

Comme beaucoup de microalgues, *Emiliana huxleyi* existe sous deux formes : une forme diploïde se reproduisant essentiellement par mitose et une forme haploïde pouvant se multiplier par mitose ou par fusion de deux cellules haploïdes pour former une cellule diploïde. Ces deux formes sont extrêmement différentes du point de vue morphologique. La forme diploïde est non mobile et couverte de coccolithophores tandis que la forme haploïde, beaucoup plus petite, ne possède pas de couche calcique mais est rendue mobile par la présence de deux flagelles (Green et al., 1996; Houdan et al., 2004).

Des chercheurs travaillant sur l'apparition et la disparition des efflorescences algales ont identifié un virus capable d'infecter *Emiliana huxleyi* et de détruire rapidement la population algale (Wilson et al., 2002), qui s'est avéré appartenir à une famille spécifique de virus à ADN double brin

marins (Schroeder et al., 2002). La grande particularité de cette relation hôte-pathogène est que ces virus ne sont capables d'infecter que la forme diploïde d'*Emiliana huxleyi*, la forme haploïde étant totalement résistante à l'infection (Frada et al., 2008). De plus, il semble qu'en présence de virus, des cellules haploïdes apparaissent dans une population initialement diploïde.

Les mécanismes moléculaires gouvernant le passage vers la forme haploïde sont inconnus à ce jour, mais la promotion de cette transition par la présence d'un agent viral et sa résistance à l'infection suggèrent un rôle potentiel des virus dans la régulation voire la mise en place de cette forme haploïde. Or, ce mécanisme est la définition première de la reproduction sexuée, puisque ce passage par une cellule haploïde permet, lors de la fusion entre deux cellules haploïdes au matériel génétique différent, de générer une cellule diploïde au génome nouveau. Si les avantages et les inconvénients évolutifs de la reproduction sexuée sont encore discutés (Lively and Morran, 2014), son mécanisme d'apparition reste pour l'instant un mystère et les données concernant *Emiliana huxleyi* suggèrent que les virus pourraient une fois de plus avoir joué un rôle essentiel dans l'évolution de leurs hôtes.

Par ailleurs, l'étude des interactions entre *Emiliana huxleyi* et les virus susceptibles de l'infecter a conduit à la conception d'un nouveau modèle d'interactions hôte-pathogène. En effet, le modèle le plus couramment utilisé est celui de la course à l'armement : la cellule met en place un facteur antiviral, le virus évolue pour contrecarrer ce facteur, la cellule évolue à nouveau pour bloquer le virus malgré cette première résistance et ainsi de suite. Ce modèle porte également le nom d'hypothèse de la Reine Rouge, par analogie avec l'œuvre de Lewis Carroll, car les deux protagonistes y évoluent en permanence sans parvenir à chambouler l'équilibre initial.

Or cette dynamique évolutive ne correspond pas à ce qui est observé pour *Emiliana huxleyi* et ses pathogènes. En effet, il s'avère que le génome de cette algue est extrêmement stable malgré les extinctions massives qu'elle subit régulièrement (Martínez Martínez et al., 2007). Cela a amené les chercheurs ayant observé la différence de susceptibilité des formes diploïde et haploïde de l'algue aux infections à proposer un nouveau modèle, baptisé la théorie du chat du Cheshire, toujours en référence à l'œuvre de Lewis Carroll. Comme ce personnage de fiction, *Emiliana huxleyi* est capable, en passant sous forme haploïde, de se rendre invisible en cas de danger pour ne redevenir diploïde qu'une fois le danger passé (Frada et al., 2008).

Ce nouveau modèle de co-évolution, observé dans le cas de l'infection d'un hôte unicellulaire, est certes difficile à imaginer dans le cas d'un hôte pluricellulaire, mais peut constituer une base de réflexion pour comprendre d'autres dynamiques d'interactions entre des communautés d'organismes, notamment les interactions proie-prédateur.

C / Découvertes de mécanismes cellulaires fondamentaux : l'export nucléaire

1/ Identification de Crm1, une karyophérine majeure

L'étude de l'export nucléaire des ARN viraux a joué un rôle fondamental dans la compréhension des mécanismes moléculaires d'export nucléaire des composants cellulaires. En effet, si les principes généraux avaient été précédemment caractérisés, des molécules clés de ces voies d'export, telles que Crm1 et TAP, demeuraient inconnues.

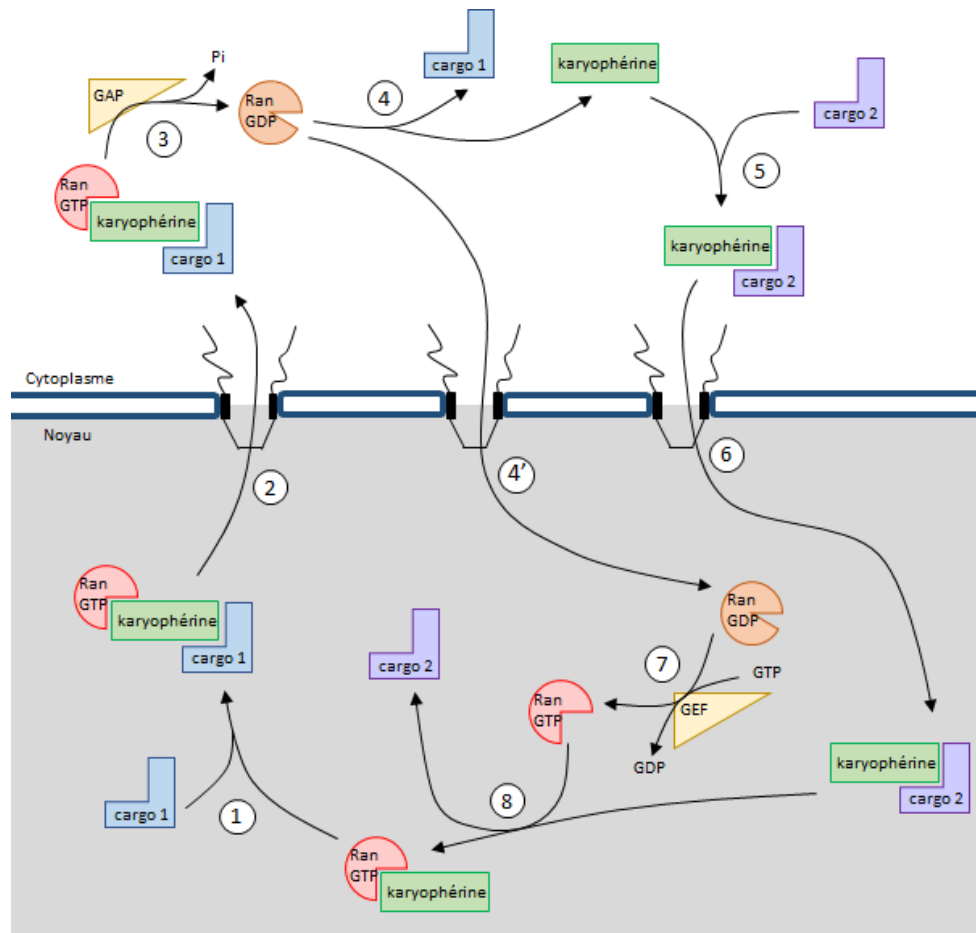


Figure 4 : schéma récapitulatif des mécanismes de transport nucléaire impliquant Ran et les karyophérines cellulaires. 1 : le complexe karyophérine-Ran-GTP interagit avec un cargo à exporter. 2 : le complexe tripartite est exporté hors du noyau via les pores nucléaires. 3 : la GAP (GTPase-Activating Protein) spécifique de Ran hydrolyse le GTP en GDP. 4 : cette hydrolyse conduit à la séparation des différents éléments du complexe. 4' : Ran-GDP peut être importé dans le noyau. 5 : la karyophérine interagit avec un cargo à importer. 6 : le complexe karyophérine/cargo entre dans le noyau. 7 : le GEF (Guanine nucleotide Exchange Factor) échange le GDP lié à Ran pour du GTP. 8 : l'interaction entre Ran-GTP et la karyophérine va séparer cette dernière de son cargo. Ce schéma rappelle un principe général. Il existe différentes karyophérines, certaines spécialisées dans l'export nucléaire, d'autres dans l'import et d'autres capables d'accomplir les deux fonctions.

L'identification de Crm1 est due à l'étude des mécanismes d'export des ARNm non épissés du VIH. A l'époque, les karyophérines et leur interaction avec la GTPase Ran pour moduler les transports nucléaires avaient été caractérisées (Moroianu and Blobel, 1995; Moroianu et al., 1995) (Figure 4). Il avait également été démontré que la protéine Rev du VIH-1 était capable d'interagir avec une séquence dite RRE (Rev-recognition element) présente sur les ARN non épissés du virus pour permettre leur export nucléaire (Michael H. Malim, Joachim Hauber, Shu-Yun Le, 1989) mais le mécanisme mis en jeu restait à caractériser.

La première découverte clé a été l'identification d'une séquence d'export nucléaire portée par la protéine Rev et essentielle pour son rôle dans l'export nucléaire. De plus, il a été montré que la voie d'export utilisée par cette protéine virale est différente de la voie d'export des ARNm cellulaires mais est en compétition avec les voies d'export des U snARN et de l'ARNr 5s (Fischer et al., 1995). Enfin, Crm1, dont le rôle dans le transport nucléaire n'était alors que putatif, a été identifié comme un partenaire de Rev, servant d'intermédiaire entre Rev et les pores nucléaires pour permettre l'export nucléaire (Neville et al., 1997). Le rôle de cette protéine de la famille des importines β dans le transport nucléaire a ainsi été caractérisé pour la première fois.

Les travaux menés depuis ont permis de mieux caractériser les fonctions de Crm1. Ainsi, cette protéine est désormais considérée comme un des pivots de l'export nucléaire. En effet, si d'autres virus que le VIH sont capables de l'utiliser pour permettre l'export de leurs ARN, notamment certains papillomavirus (Onder et al., 2015), Crm1 est aussi essentielle à l'export de différents composés cellulaires. Ainsi, Crm1 est impliquée dans la biogenèse et l'export nucléaire des ribosomes des eucaryotes (Thomas and Kutay, 2003), des U snARN (Fornerod et al., 1997) mais également de certains ARNm (Brennan et al., 2000). En plus de sa fonction dans l'export nucléaire, Crm1 a été identifiée comme jouant un rôle de coordination de différents processus au cours de la mitose, notamment la mise en place des microtubules kinétochoriens (Knauer et al., 2006), et jouerait un rôle dans la structuration de complexes nucléoplasmiques (Verheggen and Bertrand, 2012). L'ensemble des fonctions connues et supposées de Crm1 en fait un acteur majeur de la biologie nucléaire des eucaryotes (Hutten and Kehlenbach, 2007) et il est intéressant de noter que son inhibition est à l'origine de l'amélioration clinique des patients dans des essais thérapeutiques dirigés contre différents types de cancer (Kojima et al., 2013; Wang et al., 2014b; Yang et al., 2014a).

2/ Identification de TAP, acteur clé de l'export des ARNm cellulaires

L'export de la plupart des ARNm cellulaires est indépendant des karyophérines et de la GTPase Ran. Les mécanismes mis en jeu dans cette voie particulière d'export nucléaire ont également été caractérisés via l'étude de l'export nucléaire d'ARN viraux : ceux du virus Mason-Pfizer du singe (MPMV).

En effet, ce rétrovirus de type D promeut également l'export nucléaire d'ARN viraux non épissés mais ne possède pas d'équivalent de la protéine Rev du VIH. Des chercheurs ont remarqué que l'ARN du MPMV contient une séquence de 219 nucléotides capable de permettre l'export des ARN du VIH en l'absence de Rev (Bray et al., 1994). Il a été prouvé que cette séquence, baptisée CTE (Constitutive Transport Element), induit l'export nucléaire par une voie différente de celle utilisée par Rev mais interagissant avec la voie d'export des ARNm cellulaires (Saavedra et al., 1997).

L'analyse par spectrométrie de masse des protéines cellulaires capables d'interagir avec le CTE du MPMV a permis d'identifier Tap (Grüter et al., 1998), dont l'homologue Mex67p chez la levure était connu comme nécessaire à l'export des ARNm grâce à son interaction avec l'ARN et avec les pores nucléaires (Segref et al., 1997). L'implication de Tap dans l'export des ARNm et des ARN contenant un CTE a alors été caractérisée (Grüter et al., 1998). Il a depuis été montré que Tap fonctionne en dimère avec la protéine NXT1 et que ce dimère peut interagir directement avec les pores nucléaires pour permettre l'export des ARNm de façon indépendante de la GTPase Ran (Clouse et al., 2001; Lévesque et al., 2001; Wiegand et al., 2002).

Ainsi, les deux voies principales d'export des ARN cellulaires, celle dépendant de Crm1 et de Ran et celle dépendant du dimère Tap/NXT1 ont été identifiées via l'étude des mécanismes d'export d'ARN viraux. De façon intéressante, on constate que les mécanismes d'export des ARN de certains virus ne sont toujours pas élucidés aujourd'hui. C'est notamment le cas des ARN du virus du sarcome de Rous, qui possèdent une séquence du même type que le CTE du MPMV mais dont on ne connaît pas le partenaire cellulaire (Paca et al., 2000). Des découvertes sur les voies d'export des ARN grâce à l'étude des mécanismes viraux sont sans doute encore à venir.

D/ Quand les virus deviennent des outils

1/ Outils pour les chercheurs : exemple du traçage transneuronal

Les virus sont couramment utilisés en laboratoire, à la fois comme objets d'étude et comme outils car les propriétés de certains virus permettent de répondre à des questions inabordables par d'autres moyens.

En neurologie par exemple, une des problématiques clés est la compréhension de l'organisation des réseaux neuronaux. Savoir quelles parties du système nerveux communiquent entre elles et de quelles façons permet de mieux comprendre les modalités exactes de transmission de signaux que l'on peut vouloir supprimer ou restaurer pour traiter des pathologies. Il est donc important de pouvoir suivre les réseaux neuronaux à l'échelle cellulaire. Or les méthodes classiques de marquage, indépendantes des virus, restent limitées. Selon la technique choisie, elles peuvent permettre de visualiser l'ensemble des neurones activés par un stimulus mais sans mettre en évidence les connections entre ces neurones, en activant potentiellement plusieurs réseaux sans pouvoir les distinguer et en ne révélant pas les neurones non activés au sein des réseaux concernés. Des techniques plus précises existent, reposant sur le marquage d'un neurone particulier et sur la transmission de ce marquage de façon antérograde ou rétrograde aux neurones connectés avec le neurone marqué. Cependant ces techniques reposent sur des agents de marquage qui sont dilués lorsqu'ils progressent le long des réseaux neuronaux et la diminution rapide du signal permet rarement de visualiser plus de deux neurones consécutifs (Loewy, 1998). Les premiers traçages transneuraux reposant sur l'utilisation de virus neurotropes de type herpesvirus chez le rat et la souris ont donc constitué une réelle révolution en neurologie (Martin and Dolivo, 1983; Ugolini et al., 1987).

En effet, l'utilisation de virus neurotropes pour visualiser les réseaux neuronaux (par immunofluorescence ou immunohistochimie) présente deux avantages majeurs. Tout d'abord, contrairement aux autres marqueurs, les virus sont capables de se multiplier et donc d'auto-amplifier le signal qu'ils constituent après infection d'une nouvelle cellule. Ainsi, il devient possible de visualiser le marquage de plusieurs neurones consécutifs, dont le nombre varie en fonction de la durée de l'infection avant sacrifice de l'animal. Par ailleurs, il a été montré que les virus utilisés dans les expériences de traçage transneuronal ne se transmettaient que via les connexions synaptiques (Enquist, 2002; Ugolini, 1995), ce qui permet de visualiser des réseaux de neurones fonctionnellement connectés.

Deux types de virus sont utilisés pour le traçage transneuronal : des herpesvirus (dont le virus de la pseudo-rage) et des rhabdovirus. Cela vient essentiellement du fait que ces virus ont des tropismes différents : après infection intranasale chez la souris, le virus de la pseudo-rage infecte les systèmes sympathique, parasympathique et sensoriel mais pas le système olfactif alors que le virus de la stomatite vésiculaire n'infecte que le système olfactif (Sabin, 1938). Par ailleurs, si le virus de la pseudo-rage a été privilégié pour l'étude des modèles murins car il n'est pas capable d'infecter l'homme, le virus l'herpès simplex 1 (HSV-1) et les rhabdovirus ont été utilisés pour l'étude des modèles primates.

Quel que soit le type de virus utilisé, des souches permettant de retracer les réseaux neuronaux dans le sens antérograde et d'autres permettant de les retracer dans le sens rétrograde ont été caractérisées (Zemanick et al., 1991). Par ailleurs les souches généralement utilisées pour retracer les réseaux neuronaux sont des souches de virulence modérée (la souche la plus utilisée de la pseudo-rage est une souche vaccinale). Cela s'explique par le fait que les souches virulentes vont se multiplier rapidement ce qui, en plus de brouiller le signal en conduisant à l'infection d'un grand nombre de neurones en peu de temps, conduit à une mort rapide des animaux.

Il est important de souligner qu'en dehors de leur tropisme neuronal et de la possession d'une enveloppe, les herpesvirus et les rhabdovirus sont très différents. Les rhabdovirus sont des virus à ARN simple brin de polarité négative dont le génome fait 12kb alors que les herpesvirus sont des virus à

ADN double brin dont le génome fait 150kb. Cela entraîne un certain nombre de différences dans l'utilisation expérimentale de ces outils. Tout d'abord, les rhabdovirus ne possèdent qu'une seule protéine d'enveloppe, RG (rabies glycoprotein) ce qui les rend très spécifiques d'un type cellulaire voire d'une zone neuronale donnée : le virus de la rage n'est ainsi capable d'infecter les neurones qu'au niveau des terminaux présynaptiques (Lafon, 2005). En revanche les herpesvirus ont plusieurs gènes codant des protéines d'enveloppe ce qui leur permet de cibler un plus grand nombre de récepteurs (McGeoch et al., 1988).

En plus de cette spécificité différente d'entrée dans la cellule cible, la différence de taille et de complexité entre ces deux familles virales fait que la pathogenèse des rhabdovirus est mieux comprise à l'échelle cellulaire que celle des herpesvirus, mais que le génome de ces derniers est plus facilement modifiable. En effet, le développement des techniques de génie génétique a permis d'améliorer les virus utilisés dans le traçage transneuronal, en maîtrisant plus finement leur pathogénicité, en influençant leur direction de dissémination, en leur faisant exprimer des gènes rapporteurs mais surtout en ciblant plus précisément les cellules infectées.

De fait, la principale limite à l'utilisation de virus neurotropes dans le traçage transneuronal est le manque de maîtrise lors de l'infection du premier neurone dont on souhaite tracer les connexions. Si trop de neurones, proches physiquement, sont infectés simultanément, cela peut brouiller le suivi des réseaux correspondants. De grandes avancées ont pu être faites en combinant la manipulation des génomes viraux et la manipulation des génomes des modèles murins associés. Dans le cas du virus de la pseudo-rage, des souches virales ne se répliquant qu'en présence de la recombinaison Cre ont été générées. Utilisées sur des souris n'exprimant Cre que dans certains types cellulaires ou dans certaines zones du cerveau, elles permettent un suivi plus précis des réseaux neuronaux (DeFalco et al., 2001). L'organisation très compacte du génome des rhabdovirus complique sa modification, néanmoins une souche transgénique a été produite en remplaçant la protéine d'enveloppe originelle du virus par une autre protéine d'enveloppe, reconnaissant un récepteur non exprimé par les Mammifères. Utilisée en combinaison avec des souris transgéniques exprimant ce récepteur dans certains neurones, cette souche de virus de la rage permet également d'améliorer le traçage transneuronal (Wickersham et al., 2009).

2/ Outils thérapeutiques

a/ Production de vecteurs viraux

Les virus sont aussi utilisés en tant qu'outils pour leur capacité de transport d'information génétique, en laboratoire mais également en médecine, dans le cadre de thérapies géniques ou de production de vaccins.

Historiquement, l'existence de transfert de propriétés sous forme d'ADN d'une cellule à l'autre avait été démontrée chez les bactéries avant même que l'ADN ne soit identifié comme support de l'information génétique (Avery et al., 1945) et le rôle des bactériophages, virus bactériens, dans ces transferts avait été identifié assez rapidement (N. D. Zinder and J. Lederberg, 1952). Ce processus de passage d'information génétique d'une bactérie à l'autre via un virus porte le nom de transduction. L'idée que des transferts similaires pourraient avoir lieu ou pourraient être mis en place entre des cellules eucaryotes via l'utilisation d'autres virus a émergé dès les années soixante, liée à celle d'utiliser ces outils potentiels à des fins thérapeutiques (Tatum, 1966).

Il a fallu attendre la découverte des outils moléculaires appropriés, notamment les enzymes de restriction, pour que des chercheurs mettent au point la production de vecteurs viraux capables de contenir et de transmettre une information génétique donnée. Des travaux ont été menés sur différents types de virus pour aboutir à la mise au point de systèmes d'expression génétique différents. Un des premiers virus utilisé en tant que vecteur dans les cellules eucaryotes a été le polyomavirus SV40, un petit virus non enveloppé à ADN double brin. En effet, des chercheurs se sont rendu compte que ce virus était capable de produire des particules virales n'ayant encapsidé que de l'ADN cellulaire (Levine and Teresky, 1970; Trilling and Axelrod, 1970). Il a ensuite été montré que ce virus pouvait être

utilisé pour incorporer de l'ADN dans des cellules murines (Grady et al., 1970) ou humaines (Qasba and Aposhian, 1971).

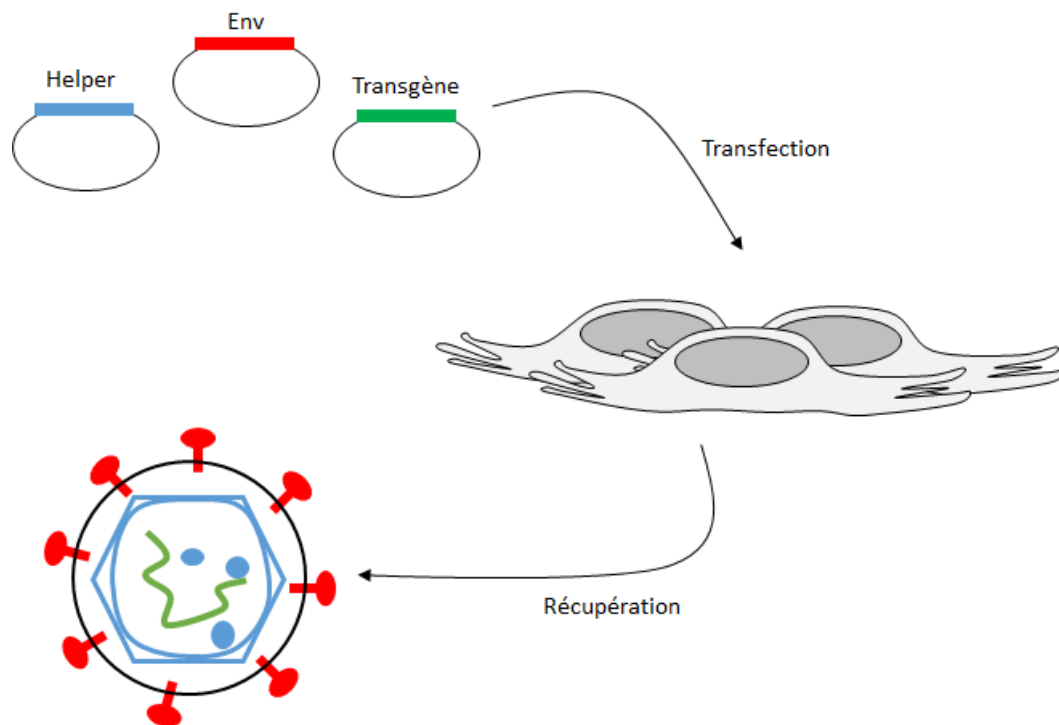


Figure 5 : principe de production de vecteurs viraux. Des plasmides codant pour le transgène d'intérêt, les gènes nécessaires à la production des particules virales (plasmide dit « helper ») et éventuellement une protéine d'enveloppe sont transfectés dans des cellules productrices. Quelques jours plus tard, des particules virales sont produites par ces cellules dans le surnageant de culture et récupérées par le manipulateur. Elles contiennent la protéine d'enveloppe choisie, les protéines structurales et les enzymes codées par le plasmide helper et du matériel génétique codant pour le transgène.

Quel que soit le type de virus utilisé, le principe de production d'un vecteur viral est toujours similaire (Figure 5). Des cellules productrices sont co-transfectées avec un plasmide contenant le transgène choisi et une séquence d'encapsidation, un plasmide dit « helper » contenant les gènes viraux nécessaires à la production des particules virales et éventuellement un troisième plasmide contenant le gène d'enveloppe virale qui permettra aux particules produites d'infecter le type cellulaire voulu. Le transgène peut contenir un ou plusieurs gènes, sous le contrôle de promoteurs choisis pour en contrôler l'expression. Quelques jours après transfection, le surnageant de culture contenant les particules virales est récupéré et ces dernières sont éventuellement purifiées par ultracentrifugation avant d'être dosées puis congelées ou utilisées pour traiter les cellules cibles (Skubis-Zegadło et al., 2013; Tiscornia et al., 2006).

Des particules virales peuvent être produites de cette manière pour sur-exprimer des gènes dans les cellules cibles (en incorporant directement les séquences d'acide nucléique correspondantes), pour sous-exprimer des gènes dans les cellules cibles (en incorporant les séquences d'ARN interférents ciblant les ARNm correspondants) ou pour introduire uniquement des protéines dans les cellules cibles (via leur interaction avec les composants des virions utilisés).

De tels outils sont utilisés pour différentes applications. Tout d'abord, en laboratoire, la production de vecteurs viraux permet la transduction de cellules potentiellement difficiles à transfecter, comme les cellules primaires, pour leur faire exprimer un transgène. L'utilisation de marqueurs de sélection, comme des gènes de résistance à des antibiotiques, et de vecteurs rétroviraux permettant l'intégration du transgène dans le génome cellulaire donnent la possibilité de sélectionner

des lignées stables contenant un transgène et l'exprimant de façon constitutive ou inducible selon le promoteur choisi (Trono, 2000).

Par ailleurs, les vecteurs viraux jouent deux rôles fondamentaux en thérapie : ils peuvent être utilisés dans le cadre de thérapies géniques pour modifier l'expression de gènes particuliers dans des cellules spécifiques ou comme vecteurs vaccinaux pour permettre la présentation d'antigènes d'agents infectieux dans un contexte non pathogène.

b/ Applications de différents types de vecteurs

Dépendamment de l'application finale, il importe de choisir le type de vecteur le plus approprié à son contexte d'utilisation (Tableau 2). En laboratoire, le but de la transduction est le plus souvent d'obtenir une expression à long terme du transgène. Pour cela, les vecteurs les plus adaptés sont ceux qui permettent l'insertion du transgène dans le génome cellulaire : les vecteurs dérivés des rétrovirus et des lentivirus.

In vivo, l'utilisation de ces vecteurs est délicate car ils sont susceptibles de perturber le fonctionnement de la cellule par mutagenèse insertionnelle et notamment de favoriser le développement de tumeurs mais ce problème ne se pose pas pour l'utilisation in vitro. En revanche, il est important de savoir si les cellules ciblées continuent à se diviser ou non. En effet, la plupart des vecteurs rétroviraux sont dépendants de la mitose et il est nécessaire d'utiliser des vecteurs lentiviraux pour insérer un transgène dans le génome d'une cellule ne se divisant pas (Trono, 2000).

Les vecteurs rétroviraux et lentiviraux n'étant capables d'infecter que certains types cellulaires, leur enveloppe est généralement remplacée par celle du Virus de la Stomatite Vésiculaire, la protéine VSVg, qui est capable de permettre l'entrée du virus dans n'importe quel type cellulaire (Burns et al., 1993; Naldini et al., 1996). On parle alors de vecteurs pseudotypés. L'autre avantage des vecteurs rétroviraux et lentiviraux est qu'ils permettent le transfert de séquences d'acide nucléique relativement longues, pouvant aller jusqu'à 10kb (Mátrai et al., 2010).

	Taille du génome (kb)	Immunogène	Persistance du génome viral	Intégration	Dépendant de la mitose
Adénovirus	37	oui	non	non	non
Aav	4,7	non	oui	rare	non
Rétrovirus	10	non	oui	oui	oui
Lentivirus	10	non	oui	oui	non

Tableau 2 : rappel des propriétés des principaux types de vecteurs viraux.

En thérapie génique, les vecteurs les plus utilisés sont ceux dérivés des adénovirus (22,5% des essais cliniques mondiaux, contre 19,2% pour les rétrovirus et 4,7% pour les lentivirus) (The Journal of Gene Medicine). Les adénovirus sont des virus à ADN double brin non enveloppés dont la capside icosaédrique mesure environ 90nm de diamètre. Leur génome d'environ 35kb est terminé par des séquences répétées appelées ITR (Inverted Terminal Repetitions). Ce virus est un pathogène humain bénin dont la majorité des infections sont asymptomatiques. Il est facile à produire in vitro et la protéine Fibre à l'origine de son tropisme cellulaire est adaptable par insertion de peptides (Capasso et al., 2014). Les vecteurs adénoviraux permettent l'expression forte et transitoire des transgènes indépendamment du processus de division cellulaire.

Les deux inconvénients majeurs des vecteurs adénoviraux sont leur forte propension à la recombinaison lors de leur production et leur caractère immunogène. La première de ces observations a conduit à la création de vecteurs appelés « gutless », littéralement vidés de l'ensemble des protéines virales originelles et ne contenant plus que les transgènes, les ITR et la séquence d'encapsidation. Cela rend hautement improbable une recombinaison avec l'ensemble des gènes du plasmide « helper » qui

conduirait à la production de virus réplicatifs. Le caractère immunogène de ces vecteurs les rend inappropriés dans certains types de thérapies (Raper et al., 2003) mais particulièrement adaptés dans les cas où une réponse du système immunitaire est souhaitée, notamment la vaccination et les thérapies géniques anticancéreuses. La première autorisation de mise sur le marché pour des thérapies géniques concerne d'ailleurs la Gendicine, un vecteur adénoviral permettant de restaurer l'expression de la protéine p53 et commercialisé en Chine pour les cancers de la tête et du cou (Pearson et al., 2004). Il est également intéressant de noter que la plupart des thérapies géniques en essais cliniques ciblent des pathologies cancéreuses (64,2% des cas) (The Journal of Gene Medicine).

Un autre type de vecteur couramment utilisé a connu plusieurs succès en thérapie génique vétérinaire : les vecteurs dérivés des virus associés aux adénovirus, ou AAV. Les AAV sont des parvoviridés du genre dependovirinae : ils ne peuvent compléter leur cycle de réplication qu'en cas de co-infection par un adénovirus et connaissent une phase de latence dans le cas contraire. Cela complique quelque peu leur production in vitro sans pour autant la rendre impossible. Les AAV sont des virus non enveloppés de 20nm de diamètre au génome ADN simple brin très court (4,7kb) dont les extrémités sont structurées en épingles à cheveux. 80% de la population humaine est séropositive vis-à-vis de l'AAV2, qui est non pathogène chez l'homme (Luo et al., 2015; Skubis-Zegadło et al., 2013). Les vecteurs dérivés des AAV permettent une expression permanente du transgène indépendante de la mitose, bien que les événements d'intégration du génome viral dans le génome cellulaire soient très rares, et peuvent être adaptés pour cibler certains types cellulaires. Ils ont ainsi permis de soigner la maladie de Hurler, une pathologie cérébrale, chez le chien (Ciron et al., 2006), l'hémophilie B chez le chien par la transduction de cellules hépatiques (actuellement en essai clinique chez l'homme) (Manno et al., 2006) et la myopathie de Duchenne chez la souris (Goyenvallé et al., 2004) et chez le chien (Vulin et al., 2012). En plus des muscles, du système nerveux central et du foie, les vecteurs dérivés des AAV peuvent cibler les poumons et la rétine (Weber et al., 2003).

Même s'ils sont moins courants et ont abouti à moins de succès thérapeutiques, des vecteurs dérivés de poxvirus (dont la vaccine) et du virus de l'herpès simplex sont également utilisés en essais cliniques (Wirth et al., 2013).

Le domaine d'application où existe la plus grande diversité dans les vecteurs utilisés est sans conteste celui de la vaccination. En effet, contrairement à la plupart des autres thérapies, les vecteurs utilisés peuvent parfois être réplicatifs et des vecteurs sont souvent utilisés pour exprimer les antigènes d'autres pathogènes, viraux ou non. En plus des lentivirus, rétrovirus, adénovirus, AAV, poxvirus et herpesvirus, il est possible d'utiliser en vaccination des picornavirus, des flavivirus et des virus à ARN de polarité négative comme ceux de la rougeole, de la rage ou de la grippe (Liu, 2010). Des études sont même en cours pour adapter un rhabdovirus de poisson à l'utilisation en tant que vaccin chez la souris (Nzonza et al., 2014).

Les critères primordiaux pour l'établissement d'un vaccin à partir d'un vecteur viral sont la rapidité d'établissement de la mémoire immunitaire (diminution du nombre d'injections), la capacité d'induire une immunité cellulaire et humorale, l'absence d'effets secondaires, un coût réduit associé à une capacité de production à grande échelle, une efficacité à long terme (diminuer les injections de rappel) et la possibilité d'être utilisé en pédiatrie. Certains virus correspondent particulièrement bien à ce cahier des charges, ce qui en fait des candidats de choix pour être utilisés comme vecteurs vaccinaux (Liu, 2010). C'est notamment le cas du virus de la rougeole, qui induit une forte réponse immunitaire, cible plusieurs types de cellules immunitaires, permet une protection à vie, peut supporter jusqu'à 5kb de transgène et ne présente pas ou peu d'effets secondaires (Lorin et al., 2005). Le seul inconvénient de ce vecteur vaccinal est qu'il doit être utilisé en pédiatrie car la plupart des adultes ont déjà été confrontés au virus sauvage de la rougeole.

II/ Compréhension des mécanismes de l'oncogenèse

A/ Evolution pathologique des caractéristiques cellulaires

1/ Définitions et concepts principaux

Le bon fonctionnement des organismes pluricellulaires repose entre autres sur le fait que chaque cellule est régulée de façon à remplir une fonction précise. Ces régulations concernent aussi bien la localisation, la prolifération, la différenciation que la mort de chacune des cellules de l'organisme. L'échappement aux mécanismes de contrôles correspondants peut entraîner de graves pathologies. Le cancer est ainsi dû à des perturbations permettant à une cellule de ne plus répondre à ces mécanismes de contrôle. Cette cellule n'est alors plus soumise à ses régulations d'origine mais répond aux pressions évolutives classiques indépendamment du reste de l'organisme. Si elle s'adapte aux conditions dans lesquelles elle est maintenue, la cellule modifiée peut se multiplier de façon incontrôlée et, à long terme, causer un cancer (Alberts et al., 2008; Gerlinger et al., 2014).

Cette définition du cancer comme l'apparition d'une masse cellulaire proliférant sans répondre aux mécanismes de contrôle de l'organisme permet de souligner qu'il existe en réalité autant de cancers que de patients. En effet, la nature du cancer dépend à la fois de la nature de la cellule initiatrice (type cellulaire et localisation dans l'organisme) et de la façon dont elle s'est affranchie des mécanismes de contrôle (Diaz-Cano, 2015).

Quel que soit le cancer, les étapes précoces au niveau cellulaire restent les mêmes (Figure 6, haut). Pour les comprendre, il est nécessaire d'avoir une vision claire des conditions classiques de prolifération des cellules au sein des organismes pluricellulaires.

Chez ces organismes, les cellules sont pour la plupart soumises à la limite de Hayflick, c'est-à-dire qu'elles ne peuvent se diviser qu'un nombre limité de fois, autour de 50 pour les cellules humaines (Hayflick, 1965). Cette limite est due au raccourcissement progressif des télomères, des séquences répétées localisées aux extrémités des chromosomes, à chaque division cellulaire (Harley et al., 1990). Les cellules souches sont parmi les seules cellules de l'organisme à ne pas être soumises à cette limitation du nombre de divisions car elles expriment une enzyme, la télomérase, qui est capable de rallonger les télomères (Batista, 2014; Greider, 1990). En plus de cette limitation intrinsèque de la prolifération cellulaire, les cellules saines sont soumises à une inhibition de contact. Autrement dit, lorsqu'au sein d'un tissu la surface de chaque cellule est intégralement en contact avec la surface d'autres cellules, les cellules ne se divisent plus (Abercrombie, 1967). La prolifération cellulaire n'est possible que lorsqu'elle est rendue nécessaire pour le renouvellement de cellules préalablement détruites, par des lésions ou par apoptose.

Dans ce contexte, la cellule à l'origine du cancer, ou cellule initiatrice, subit une modification de son transcriptome et devient capable de proliférer indéfiniment sans être limitée par le raccourcissement des télomères ni par l'inhibition de contact. Elle est dite transformée. Les cellules transformées sont fréquemment utilisées en laboratoire sous la forme de lignées cellulaires, leur capacité de prolifération permettant de les maintenir en culture pendant plusieurs semaines (Macpherson, 1970).

Il est important de souligner que si l'ensemble des cellules cancéreuses sont transformées, toutes les cellules transformées ne sont pas cancéreuses (Gee and Harris, 1979). En effet, une cellule est dite cancéreuse lorsqu'elle peut entraîner le développement d'un cancer. Or ce processus requiert d'autres particularités que la capacité de proliférer sans limite. Historiquement, le caractère cancéreux des cellules était évalué en les injectant à des souris et en observant le développement de tumeurs chez ces animaux. Il a ainsi été clairement montré que les cellules transformées n'induisent pas toutes le développement de tumeurs et ne sont donc pas toutes cancéreuses (Alberts et al., 2008).

Cette nomenclature à l'échelle cellulaire illustre l'évolution progressive de la cellule saine à la cellule cancéreuse. Ce principe est également applicable aux lésions causées, au niveau du tissu et de

l'organisme, par ces cellules (Figure 6, bas). Ainsi, des altérations légères de certains organes perturbant leur fonctionnement sont considérées comme des lésions pré-tumorales. Elles sont généralement traitées de façon symptomatique et n'évoluent que rarement en véritables cancers (Alberts et al., 2008). Très variables selon les organes et les tissus, ces lésions pré-tumorales ne seront pas détaillées dans le présent manuscrit.

Les lésions pré-tumorales peuvent évoluer pour devenir des tumeurs. Les tumeurs, ou néoplasmes, sont des ensembles de cellules proliférant de façon anormale et ne partageant pas l'organisation du tissu environnant. Ces tumeurs sont dites bénignes si elles ne possèdent pas de capacité invasive. Elles peuvent alors être traitées par ablation chirurgicale ou traitement local. En revanche, elles sont dites malignes lorsqu'elles sont capables d'envahir le tissu environnant voire des points distants de l'organisme via la circulation sanguine. Cette dissémination peut causer l'apparition de tumeurs secondaires, dites métastatiques, qui rendent le cancer d'autant plus difficile à traiter et aggravent le pronostic vital des patients (Alberts et al., 2008).

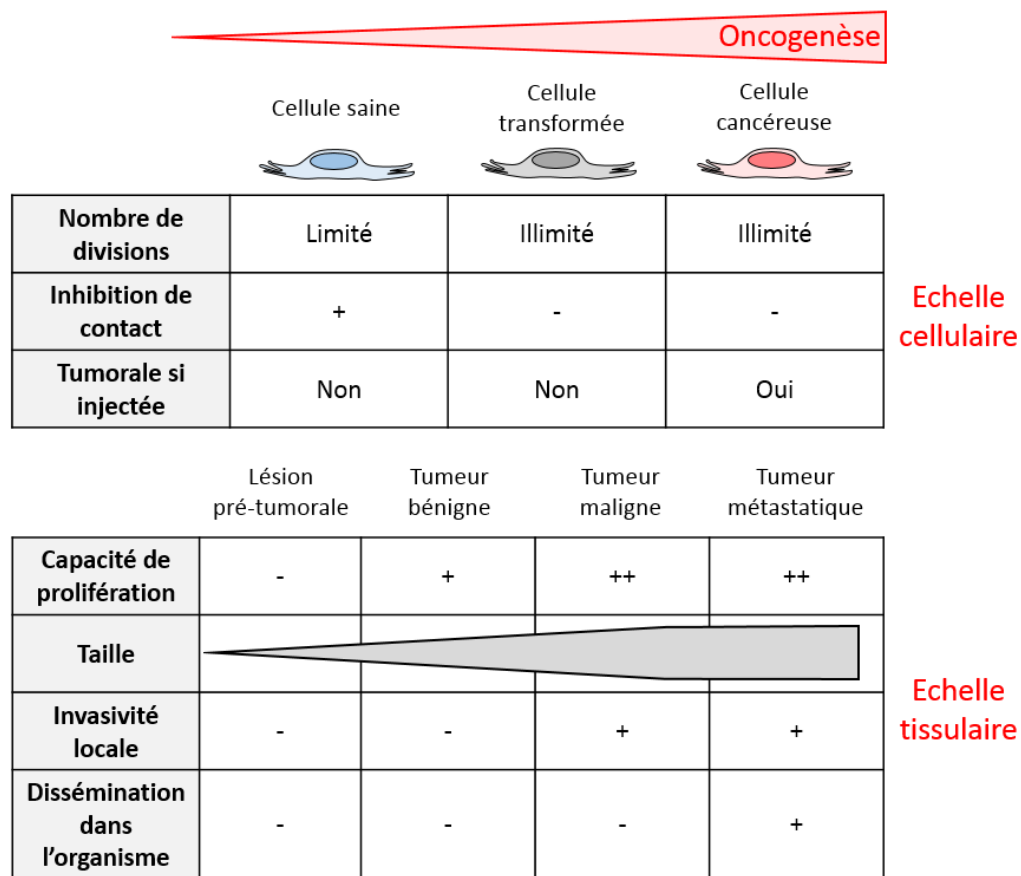


Figure 6 : récapitulatif de l'évolution des caractéristiques cellulaires et tissulaires au cours de l'oncogenèse.

Concernant l'évolution du cancer au niveau tissulaire, il est important de distinguer les tumeurs solides des tumeurs non solides. Celles-ci incluent notamment les leucémies et les lymphomes, issus de la prolifération anormale de leucocytes ou de leurs précurseurs. En effet, les cellules cancéreuses étant transportées dans la circulation sanguine, il n'est pas possible d'utiliser le critère d'invasivité pour distinguer les tumeurs bénignes des tumeurs malignes. La différence est alors faite sur la rapidité d'évolution de la pathologie (leucémie aiguë ou chronique, lymphome indolent ou agressif).

Quels que soient le type cellulaire et le tissu concernés, le cancer est en réalité un continuum d'états cellulaires et tissulaires de gravité croissante. Selon les cas, la progression vers les états les plus graves peut prendre quelques semaines ou plusieurs années, voire ne jamais avoir lieu. Ainsi, la caractérisation d'un cancer, notamment lors d'un diagnostic, nécessite d'évaluer sa probabilité d'évolution.

2/ De la cellule saine à la cellule cancéreuse

a/ Le cancer, une accumulation de dysfonctionnements

L'aggravation progressive des comportements cellulaires et des lésions vers le cancer découle du fait que cette maladie est causée par l'accumulation au fil du temps de différents dysfonctionnements. Ces dysfonctionnements ont pour origine des modifications de l'expression des gènes dans les cellules initiatrices de la maladie. Celles-ci peuvent être dues à des mutations modifiant directement la séquence codante d'un gène en entraînant des gains ou des pertes de fonctions, à des mutations des séquences régulatrices modifiant la quantité de protéine produite à partir d'un gène donné mais aussi à des changements de l'état épigénétique (Jones and Baylin, 2002) d'un gène ou d'une séquence régulatrice.

Ces changements dans l'expression des gènes, qu'ils soient dus à des mutations ou à des modifications de l'état épigénétique de séquences données, induisent une évolution du comportement des cellules concernées. L'accumulation différentielle de perturbations dans les cellules cancéreuses est couplée à un processus de sélection naturelle à l'échelle cellulaire (Figure 7). Les cellules les plus aptes à se multiplier et les plus capables d'échapper aux mécanismes de contrôle de l'organisme vont être progressivement sélectionnées et contribuer à l'aggravation de la maladie (Gerlinger et al., 2014).

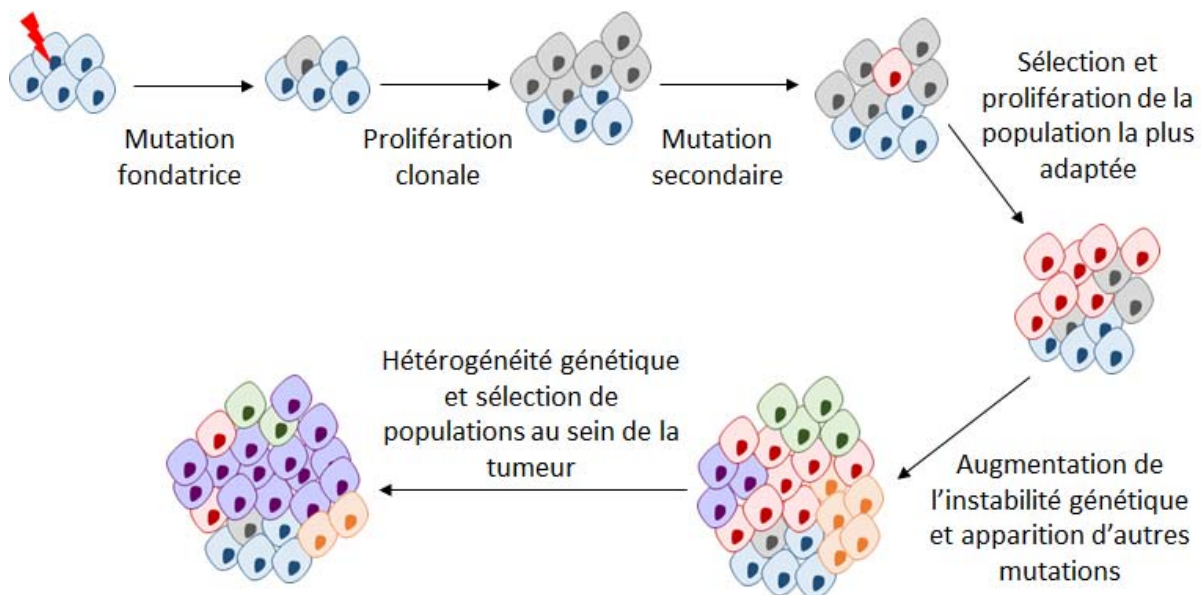


Figure 7 : représentation schématique des processus évolutifs mis en jeu au sein des tumeurs

Au-delà de cette corrélation entre accumulation de mutations et aggravation du phénotype, différentes données montrent que le cancer résulte d'une série d'événements et non d'un unique dysfonctionnement originel.

Tout d'abord, l'ADN polymérase catalysant la réplication de l'ADN au cours du cycle cellulaire possède un taux d'erreur faible mais non nul, de l'ordre de 10^{-6} mutations par gène par division cellulaire chez l'Homme. Sachant qu'environ 10^{16} divisions cellulaires ont lieu au cours de la vie d'un Homme, chacun de nos gènes a statistiquement 10^{10} occasions indépendantes d'être muté (Alberts et

al., 2008). Même en considérant que la plupart de ces mutations seraient silencieuses ou n'indui-raient pas un comportement de type cancéreux de la cellule concernée, les cancers seraient extrêmement fréquents s'ils pouvaient être causés par une seule mutation dans un gène donné.

De plus, si une seule mutation pouvait transformer une cellule saine en cellule cancéreuse, la probabilité de développer un cancer ne serait pas dépendante de l'âge. En effet, en l'absence d'agents mutagènes, la probabilité qu'une mutation apparaisse dans un gène donné ne varie pas au cours de la vie d'un individu. Un enfant et une personne âgée auraient donc la même probabilité d'avoir un cancer. Or, dans les faits, l'incidence des cancers augmente avec l'âge (Balducci and Ershler, 2005). Cette observation a conduit les scientifiques à supposer que le cancer découlait d'une accumulation de mutations avant même l'identification des gènes concernés.

A l'heure actuelle, un certain nombre de gènes impliqués dans le développement du cancer ont été caractérisés. Il est donc possible d'étudier leur séquence dans les cellules cancéreuses. Ce séquençage a permis de montrer que des mutations dans plusieurs gènes clés sont effectivement présentes dans les cellules pathologiques (Alberts et al., 2008; Gray and Collins, 2000). Par ailleurs, des expériences réalisées chez la souris ont confirmé qu'une seule mutation ne suffit pas à causer un cancer. En effet, des souris transgéniques portant une mutation impliquée dans le cancer ne développent des tumeurs qu'après un certain temps et uniquement à partir de certaines cellules (Alberts et al., 2008; van Miltenburg and Jonkers, 2012). Ceci montre que d'autres événements en plus de la mutation originelle sont nécessaires à la progression de la maladie.

b/ Caractéristiques des cellules cancéreuses et hétérogénéité tumorale

Quel que soit le type de cancer, les cellules cancéreuses partagent un certain nombre de caractéristiques connues comme les « marques de fabrique » du cancer (« Hallmarks of cancer »). La progression des connaissances sur les mécanismes d'établissement de cette pathologie et le comportement des cellules qui la causent va de pair avec une augmentation du nombre de ces caractéristiques.

Lors de leur première description, ces caractéristiques étaient au nombre de six (Hanahan and Weinberg, 2000):

- La capacité illimitée de réplication
- L'indépendance vis-à-vis des signaux de croissance
- La résistance aux inhibiteurs de croissance
- L'échappement à la mort cellulaire programmée par apoptose
- La capacité d'invasion et de métastase
- L'induction de l'angiogenèse

Les trois premières caractéristiques se retrouvent chez les cellules transformées et chez les cellules cancéreuses alors que les trois dernières sont spécifiques des cellules cancéreuses. Ces propriétés peuvent être définies au niveau moléculaire, au niveau cellulaire et au niveau de la tumeur (Figure 8, haut).

Depuis cette première description, d'autres caractéristiques partagées par les cellules cancéreuses ont été identifiées (Hanahan and Weinberg, 2011) (Figure 8, bas). Il s'agit notamment de :

- La résistance vis-à-vis du système immunitaire
- La stimulation par l'inflammation
- L'instabilité génétique
- Un changement du métabolisme cellulaire (augmentation de la glycolyse couplée à la fermentation lactique et non à la respiration mitochondriale, ou effet Warburg)

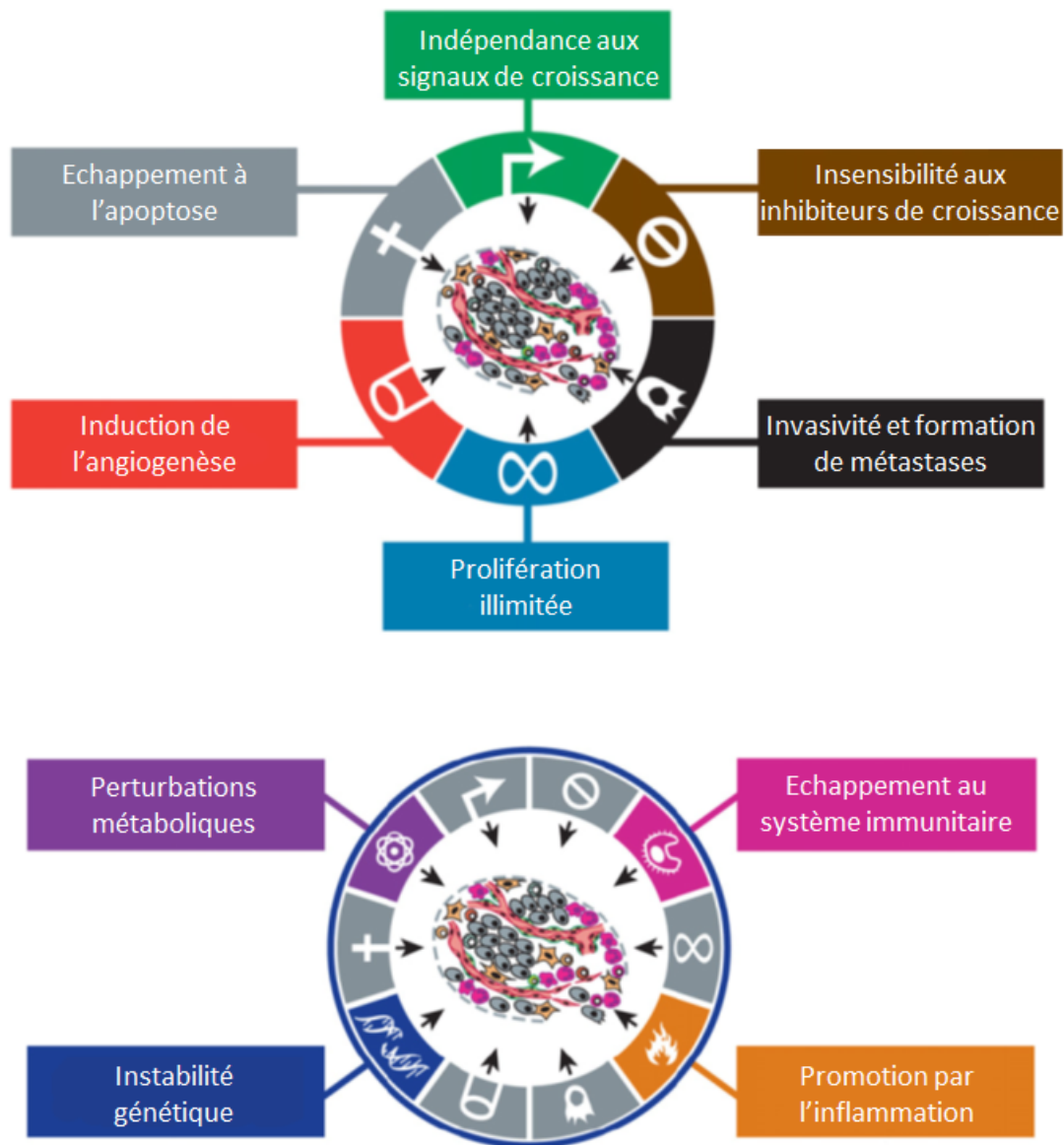


Figure 8 : récapitulatif des caractéristiques du cancer, d'après Hanahan et Weinberg (Hanahan and Weinberg, 2011).

L'acquisition progressive ou parallèle de ces caractéristiques par des cellules initialement saines conduit à leur évolution vers le cancer. Il est cependant important de préciser que toutes ces propriétés ne sont pas partagées par l'ensemble des cellules tumorales. En effet, si la tumeur est initialement causée par la prolifération anormale d'une cellule mutée et a donc une constitution originelle clonale, cette clonalité disparaît rapidement.

Les mécanismes précis conduisant à l'apparition de différents types de cellules au sein de la tumeur sont encore mal connus, et les types de cellules en question sont progressivement caractérisés (Meacham and Morrison, 2013). Néanmoins il est clair que l'instabilité génétique des cellules cancéreuses favorise l'apparition de mutations différentes dans chacune des cellules tumorales. Les cellules peuvent alors évoluer de façons différentes au sein de la tumeur (Gerlinger et al., 2014).

Parmi les différentes cellules cancéreuses, les cellules souches cancéreuses sont aujourd'hui intensivement étudiées. Contrairement aux autres cellules cancéreuses, les cellules souches cancéreuses se divisent relativement lentement mais pourraient être à l'origine directe de certains cancers. Elles seraient également particulièrement importantes pour la formation des métastases et pourraient être à l'origine d'un grand nombre de récives après traitement, leur faible taux de division

les rendant résistantes à la plupart des agents anti-cancéreux (Islam et al., 2015; Meacham and Morrison, 2013).

Enfin, il ne faut pas oublier que les tumeurs ne sont pas constituées que de cellules cancéreuses. Elles contiennent aussi des cellules saines, dites stromales, telles que des fibroblastes, des cellules endothéliales ou des leucocytes. Ces cellules ne partagent pas les caractéristiques des cellules cancéreuses mais elles favorisent le maintien et la croissance de la tumeur en lui apportant différents facteurs (oxygène transporté via le sang, sécrétion d'enzymes altérant la matrice extracellulaire...) (Bhome et al., 2015; Pietras and Östman, 2010).

B/ Comment identifier les gènes impliqués dans l'oncogénèse ?

1/ Oncogènes et suppresseurs de tumeurs

Les gènes dont la perturbation favorise le développement de cancers peuvent être classés en deux catégories principales : les oncogènes et les suppresseurs de tumeurs (Alberts et al., 2008; Klein, 1988).

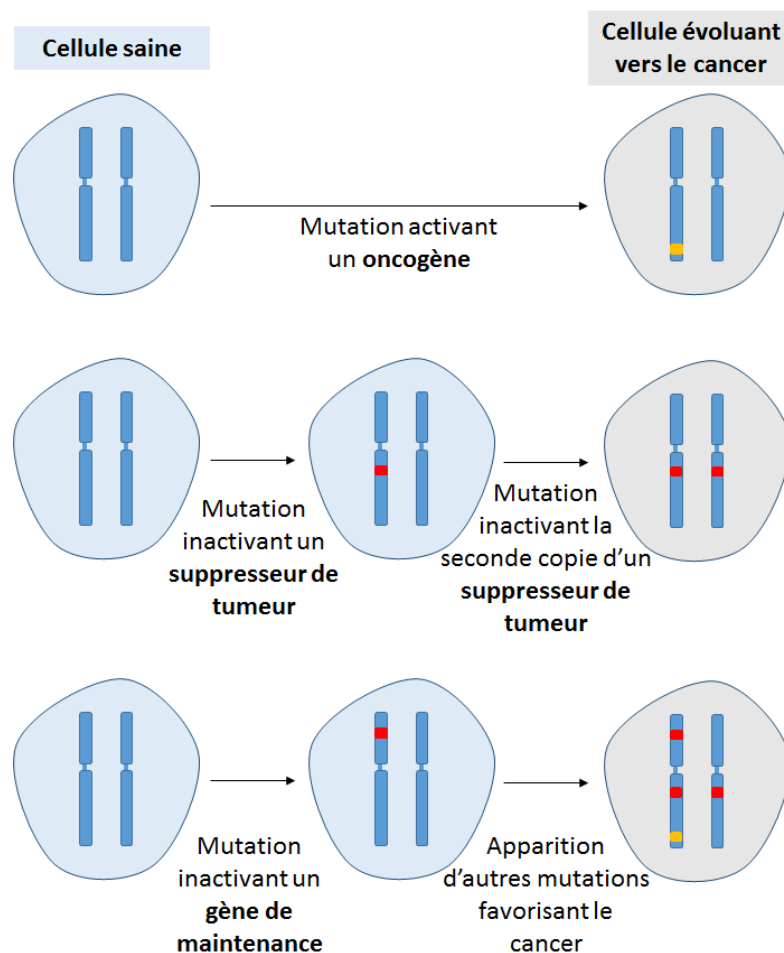


Figure 9 : schéma des trois grandes catégories de mutations favorisant le développement du cancer. Les deux chromosomes représentés correspondent aux deux chromosomes d'une même paire. Les mutations gain de fonction sont indiquées en jaune, les mutations perte de fonction en rouge.

Les oncogènes sont des gènes favorisant l'apparition des cancers (Figure 9, haut). Ils sont présents dans le génome sous forme de proto-oncogènes. Leur caractère pathologique est dû à une activation constitutive de la protéine qu'ils codent. Celle-ci peut être causée par des modifications qualitatives (mutations) ou quantitatives (duplication, surexpression...) du gène ou de ses produits (ARNm et protéine) aboutissant à un gain de fonction. L'activation d'un oncogène cellulaire est

dominante, la modification d'un seul des deux allèles cellulaires suffisant à favoriser le développement cancéreux (Murray et al., 1981).

Les suppresseurs de tumeurs, comme leur nom l'indique, ont pour fonction de limiter l'oncogenèse (Figure 9, milieu). Ils sont normalement exprimés dans les cellules et c'est cette fois-ci leur perte qui est à l'origine de l'apparition des cancers. Cette perte peut être due à des mutations ponctuelles inactivant le gène ou déstabilisant la protéine, mais aussi à des modifications épigénétiques empêchant l'expression du gène ou à des événements de recombinaison conduisant à la perte totale du gène. Ces pertes de fonction sont généralement récessives et les suppresseurs de tumeurs ne perdent leur activité que lorsque les deux allèles du gène ont été altérés (Figure 10) (Levine, 1993).

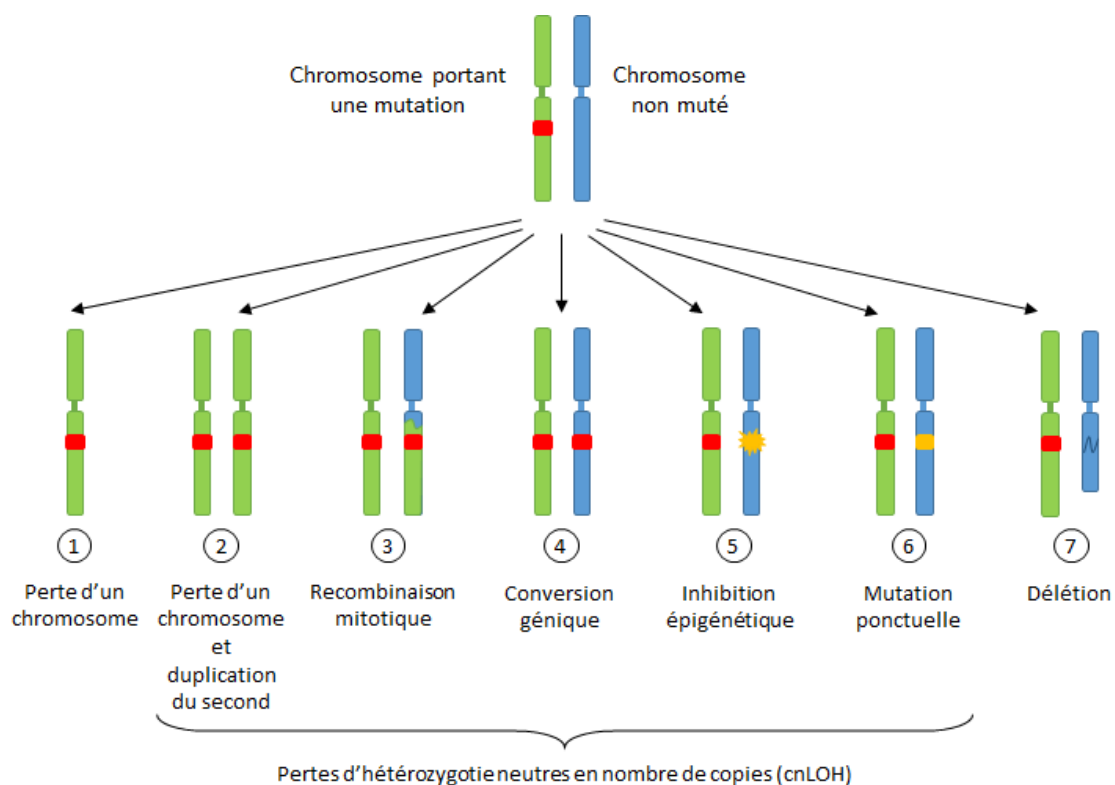


Figure 10 : schéma de sept mécanismes pouvant conduire à une perte d'hétérozygotie pour un gène donné. Les deux chromosomes représentés correspondent aux deux chromosomes d'une même paire. Si la mutation indiquée en rouge inactive un gène suppresseur de tumeur, ces mécanismes sont autant de façon d'inactiver la deuxième copie de ce gène.

A ces deux classes de gènes impliqués dans l'oncogenèse peut être ajoutée une troisième catégorie, celle des gènes de maintenance (Figure 9, bas). L'action de ces gènes sur le développement du cancer est plus indirecte que pour les précédents mais reste néanmoins essentielle. En effet, cette catégorie regroupe l'ensemble des gènes impliqués dans le maintien de la stabilité du génome. Lorsque ceux-ci sont altérés, ce qui est fréquemment le cas dans les cellules cancéreuses, la probabilité de mutation augmente. Cumulée à l'augmentation de la vitesse de prolifération des cellules, c'est-à-dire de la fréquence de réplication de leur génome, celle-ci favorise le dérèglement des oncogènes et des gènes suppresseurs de tumeurs, donc le développement du cancer (Fishel and Kolodner, 1995; Fox et al., 2013).

L'identification des proto-oncogènes et des suppresseurs de tumeur est facilitée par le développement des techniques de séquençage haut débit de génomes entiers. Néanmoins, l'instabilité génétique globale des cellules cancéreuses rend impossible l'identification de gènes impliqués dans le

cancer en comparant simplement le génome d'une cellule cancéreuse avec celui d'une cellule saine du même patient : les différences sont trop nombreuses. De plus l'expression des gènes impliqués dans l'oncogenèse peut être altérée via des modifications épigénétiques non détectables par séquençage (Jones and Baylin, 2002). Des analyses de séquençage portant sur un plus grand nombre de patients ou des méthodes différentes sont donc nécessaires.

Enfin, il faut souligner que le terme de « gène » employé pour parler des gènes impliqués dans le cancer est en réalité un abus de langage. Il serait plus approprié de parler de « séquences » d'ADN. En effet, certains microARNs sont impliqués dans l'oncogenèse, en tant qu'oncogènes ou que suppresseurs de tumeurs (Suzuki et al., 2012).

2/ Apports de l'étude des cancers familiaux et héréditaires

Bien que relativement rares, les cancers familiaux et héréditaires ont permis d'identifier plusieurs gènes impliqués dans l'oncogenèse. En effet, après caractérisation dans le cas particulier d'un cancer familial donné, plusieurs gènes se sont ensuite avérés impliqués dans le développement de nombreux cancers sporadiques.

Ainsi l'étude d'un cancer de la rétine, le rétinoblastome, a permis d'identifier le gène *Rb-1* comme étant un suppresseur de tumeur. Le rétinoblastome est un cancer rare ayant la particularité d'être causé par un nombre très réduit de mutations et touchant généralement les enfants de moins de cinq ans. Ce cancer peut être transmis de façon héréditaire ou se développer de façon sporadique. Dans la forme héréditaire, plusieurs tumeurs simultanées sont souvent observées, touchant généralement les deux yeux alors que la forme sporadique conduit à l'apparition d'une tumeur unique dans un seul œil (Cowell et al., 1988).

Chez certains patients, les cellules du rétinoblastome possèdent un caryotype anormal présentant une délétion au niveau du chromosome 13. L'étude du fragment sujet à délétion a permis de caractériser le gène *Rb-1* (Friend et al., 1986; Vogel, 1979) et de montrer que, chez la quasi-totalité des patients présentant un rétinoblastome, les deux allèles de ce gène sont non fonctionnels (Dunn et al., 1988). *Rb-1* est un suppresseur de tumeur qui a depuis été identifié comme muté dans d'autres types de cancer (JugeMourineau et al., 1997; Sauerbrey et al., 1998).

Le même type de démarche expérimentale a permis de caractériser le suppresseur de tumeur *APC* (adenomatous polyposis coli) comme agent causal de la polypose adénomateuse familiale, un type de cancer du côlon (Grodin et al., 1991; Kinzler et al., 1991). Ce gène est également muté dans certains cancers sporadiques du côlon, de l'estomac ou du pancréas (Nakamura, 1995). Les gènes de susceptibilité aux cancers du sein et des ovaires *BRCA-1* et *BRCA-2* ont également été identifiés à partir de l'étude de formes familiales du cancer du sein (Miki et al., 1994; Wooster et al., 1995).

3/ Importance historique des virus dans la compréhension de l'oncogenèse

Plus de 15% des cancers humains sont la conséquence plus ou moins directe d'une infection par un virus, alors dit oncogène (Butel and Fan, 2012). Ces virus oncogènes sont tous des virus à ADN ou des rétrovirus, à l'exception du Virus de l'Hépatite C (VHC) qui est un virus à ARN simple brin de polarité positive. Le VHC, ainsi que le Virus de l'Hépatite B (VHB) est associé aux cancers du foie. Parmi les virus à ADN, on peut notamment citer les papillomavirus, associés aux cancers du col de l'utérus et les herpesvirus. Les rétrovirus quant à eux sont associés à des leucémies (Butel and Fan, 2012).

Si les virus à ADN induisent le développement du cancer en exprimant des protéines inhibant des suppresseurs de tumeurs cellulaires, deux mécanismes différents sont possibles dans le cas des rétrovirus (Alberts et al., 2008). Cette partie revient sur ces mécanismes, qui ont permis la découverte de proto-oncogènes cellulaires.

Les rétrovirus ont la capacité de rétrotranscrire leur génome d'une forme ARN à une forme ADN et de conduire à l'intégration de cet ADN viral au sein de l'ADN cellulaire. Ce mécanisme d'insertion est susceptible de perturber le fonctionnement des cellules concernées, tout comme celui des éléments mobiles du génome, qu'ils soient d'origine rétrovirale ou non. C'est ainsi que l'étude des rétrovirus aviaires au cours du vingtième siècle a été à l'origine de la caractérisation de plusieurs gènes cellulaires impliqués dans la régulation de la tumorigenèse. La démarche expérimentale étant comparable, un seul exemple sera développé ici : celui de la caractérisation du proto-oncogène *myc*.

L'idée selon laquelle les virus peuvent être associés au développement de cancers remonte à plus d'un siècle. Ainsi, dès 1908, Ellerman et Bang ont établi un lien de causalité entre l'infection de poulets par le virus de la leucémie aviaire (VLA) et le développement de leucémies (Ellermann and Bang, 1909). La fréquence d'association était relativement faible puisque seuls 5% des animaux infectés par le virus présentaient des leucémies six mois post-infection.

En parallèle, Rous, qui travaillait sur d'autres tumeurs aviaires, a montré que celles-ci n'étaient transmissibles qu'entre des animaux de la même race et que cette transmission était indépendante de la présence de cellules tumorales dans l'inoculat (Rous, 1911). Cela l'a mené à formuler deux hypothèses : ces sarcomes pouvaient être causés par des agents infectieux ou par molécules sécrétées par les cellules tumorales. Il s'est rapidement avéré qu'un virus était à l'origine du développement de ces sarcomes (Rous and Murphy, 1912), qui a été baptisé virus du sarcome de Rous.

Ces deux exemples ont joué un rôle fondamental dans l'acceptation du rôle de certains virus dans le processus de tumorigenèse. Les rétrovirus oncogènes aviaires étudiés à l'époque pouvaient être classés dans deux catégories : certains avaient un fort pouvoir cancérigène (la quasi-totalité des animaux infectés développaient des tumeurs) alors que d'autres, comme le VLA, n'induisaient des cancers que chez une partie des animaux infectés.

Il a fallu attendre quelques décennies et le développement d'outils moléculaires, notamment le séquençage, pour comprendre l'origine de cette différence de pathogénicité. Les virus à fort pouvoir cancérigène, parmi lesquels la souche MC29 du virus de la myélocytomatose aviaire, sont retrouvés dans l'ensemble des cellules tumorales des animaux infectés mais sont absents des cellules non tumorales. L'étude du génome du rétrovirus MC29 a permis de montrer que celui-ci, pourtant considéré comme déficient car incapable de se multiplier en l'absence de co-infection par un autre rétrovirus, possédait un gène capable d'induire la transformation des cellules infectées en cellules tumorales quel que soit son site d'insertion dans le génome (Lautenberger et al., 1981). Cet oncogène viral a été baptisé *v-myc* (Figure 11, partie de gauche).

Parallèlement, des chercheurs qui travaillaient sur l'oncogenèse induite par le VLA se sont aperçu que ce virus ne possédait pas d'oncogène propre. De plus, malgré le caractère apparemment aléatoire du site d'insertion du génome viral dans le génome cellulaire, les tumeurs induites par le VLA sont des tumeurs clonales provenant de la multiplication d'une cellule où le génome viral s'est inséré dans un locus particulier. Les LTR viraux étant des séquences promotrices, l'hypothèse dominante était celle d'une oncogenèse par mutation insertionnelle : l'insertion d'un nouveau promoteur dans le génome cellulaire à un endroit précis conduirait à la surexpression d'un gène cellulaire à l'origine du développement tumoral. Cette hypothèse était renforcée par l'observation, dans les tumeurs VLA-dépendantes, d'ARN mixtes contenant des séquences virales issues du LTR accolées à des séquences d'origine cellulaire. En hybridant ces ARN avec des sondes spécifiques de différents oncogènes connus, des chercheurs ont pu montrer que, dans 85% des cas, les lymphomes causés par le VLA étaient dus à une activation par le LTR viral de l'expression d'un gène cellulaire très proche du gène *v-myc* du virus MC29 et qui a donc été appelé *c-myc* (Hayward et al., 1981) (Figure 11, partie de droite).

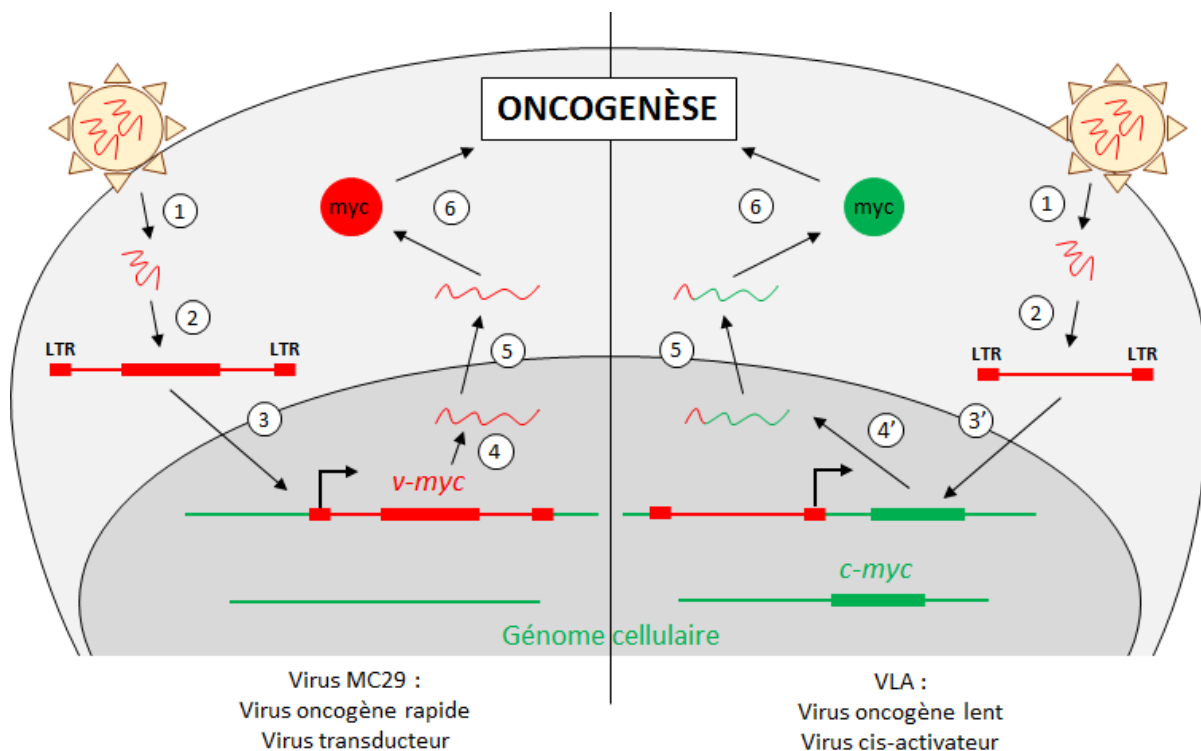


Figure 11 : schéma récapitulant les deux principaux types d'oncogenèse médiés par les rétrovirus et impliquant l'expression d'un oncogène. A gauche le cas d'un virus oncogène rapide, dit transducteur, dont le modèle classique est celui du virus du sarcome de Rous, à droite le cas d'un virus oncogène lent, dit cis-activateur, dont le modèle classique est le virus de la leucémie aviaire. 1 : infection de la cellule cible et entrée de l'ARN viral. 2 : transcription inverse pour former l'ADN viral. 3 : insertion de l'ADN viral dans l'ADN cellulaire. 3' : insertion de l'ADN viral dans l'ADN cellulaire en amont du locus *c-myc*. 4 : transcription de l'oncogène viral *v-myc* sous le contrôle du promoteur LTR viral. 4' : transcription de l'oncogène cellulaire *c-myc* sous le contrôle du promoteur LTR viral. 5 : export nucléaire de l'ARNm *myc*. 6 : traduction de la protéine *myc* et induction de l'oncogenèse.

Pour la première fois, le potentiel oncogénique d'un gène cellulaire a été identifié et il a été démontré qu'un virus oncogène pouvait agir en modifiant l'activité d'un proto-oncogène cellulaire. Les deux mécanismes décrits ci-dessus (oncogenèse par l'expression d'un oncogène viral ou par la surexpression d'un oncogène cellulaire) ont depuis été identifiés dans de nombreux cas de cancer induit par des rétrovirus.

Le seul rétrovirus oncogène n'appartenant ni à la catégorie des virus transducteurs ni à la catégorie des virus cis-activateurs est le virus HTLV-1 (virus T-lymphotropique humain). Ce virus est le premier virus dont le caractère oncogène a été formellement caractérisé chez l'homme et il agit par trans-activation de multiples gènes cellulaires favorisant le développement de leucémies (Zane and Jeang, 2014).

C/ Processus cellulaires perturbés lors de l'oncogenèse

Comme rappelé précédemment, de nombreux processus cellulaires sont perturbés lors de l'oncogenèse. Dans cette partie, seuls ceux qui sont bien connus et sont impliqués dans le développement des leucémies seront détaillés. On ne reviendra notamment pas sur la modification de la sensibilité aux activateurs et inhibiteurs de croissance, la capacité d'invasion et de métastase, l'angiogenèse, les changements métaboliques et le détournement de la réponse immunitaire.

1/ Résistance à l'apoptose

A l'échelle de l'organisme et des tissus, plusieurs mécanismes permettent en temps normal d'éliminer les cellules présentant des dysfonctionnements. Parmi ces mécanismes, l'apoptose présente la double caractéristique d'être génétiquement programmée au sein de chaque cellule et de ne pas déclencher de réaction inflammatoire. L'apoptose joue un rôle essentiel au cours du développement embryonnaire et de la vie de l'organisme. Elle est par exemple impliquée dans la morphogenèse des doigts et le renouvellement de différents tissus, dont l'épithélium intestinal. Le mécanisme d'apoptose est également déclenché lorsqu'une cellule subit des dommages à l'ADN importants (Flusberg and Sorger, 2015). Il permet donc d'éviter le maintien dans l'organisme de cellules porteuses de mutation et susceptible de dériver vers un phénotype cancéreux. L'ensemble des cancers effectivement observés découle donc de cellules ayant acquis la capacité d'échapper à l'apoptose.

L'apoptose d'une cellule peut être induite par deux voies de signalisation différentes. La voie extrinsèque dépend d'un signal extracellulaire tandis que la voie intrinsèque est déclenchée spontanément par la cellule elle-même. Ces deux voies convergent vers l'activation des caspases pro-apoptotiques (Alberts et al., 2008). Ces protéases sont présentes dans le cytoplasme sous forme de précurseurs inactifs, les pro-caspases, qui sont activés par clivage et dimérisation (Parrish et al., 2013). L'activation de ces caspases par les voies extrinsèque et intrinsèque de l'apoptose induit le clivage de différentes protéines cellulaires et l'activation d'une nucléase capable de cliver l'ADN internucléosomique, l'apoptose étant associée à un clivage de l'ADN cellulaire (Widtak, 2000). L'ensemble de ces mécanismes régulés de dégradation aboutit à la mort cellulaire par apoptose (Figure 12).

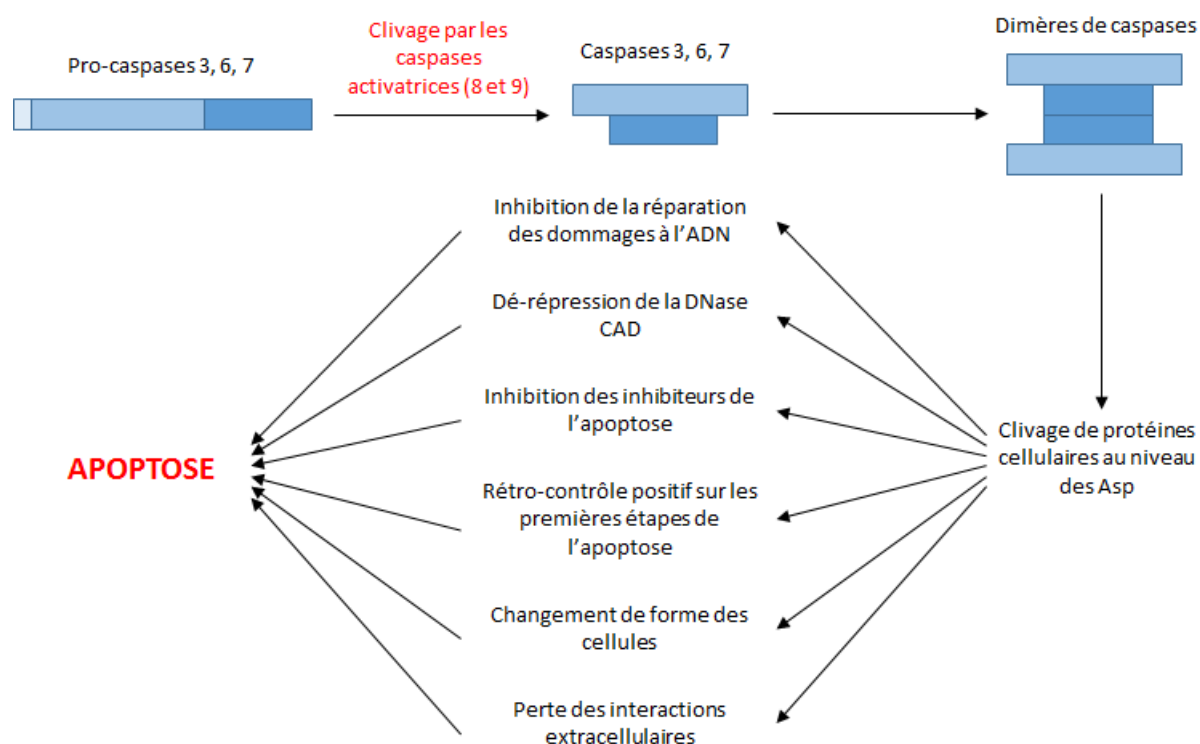


Figure 12 : schéma du mécanisme de déclenchement de l'apoptose par les caspases activatrices.

La voie extrinsèque de l'apoptose (Figure 13) est déclenchée par la reconnaissance d'un signal extracellulaire par un récepteur membranaire couplé à un domaine de mort (DD). Ce signal extracellulaire peut-être localisé à la surface d'une autre cellule (ligand Fas) ou circulant (TNF α et β).

L'interaction du récepteur avec le ligand extracellulaire va induire sa trimérisation, permettant l'interaction de ses DD intracellulaires avec les DD de protéines adaptatrices telles que FADD (protéine associée à Fas avec Domaines de mort). Ces protéines adaptatrices recrutent et activent la caspase 8 qui déclenche ensuite une cascade d'activation des autres caspases conduisant à l'apoptose (Verbrugge et al., 2010).

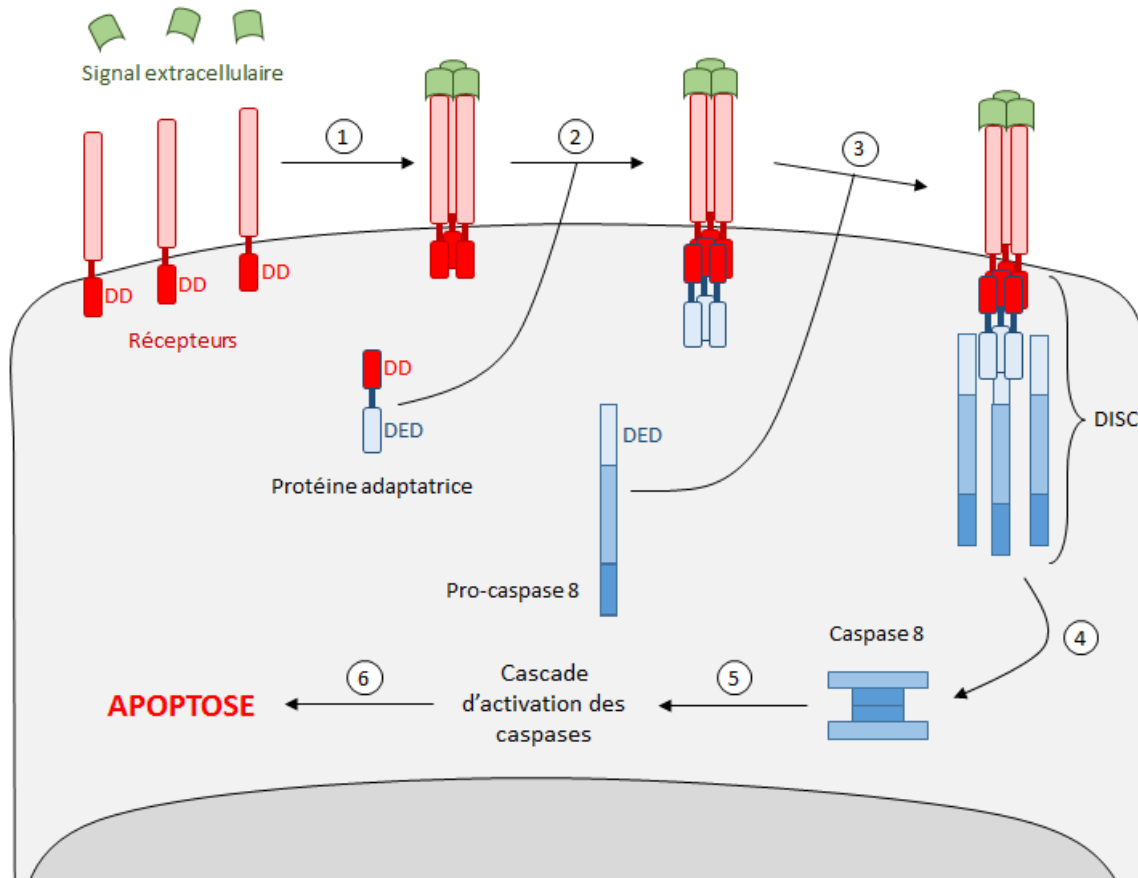


Figure 13 : représentation schématique de la voie extrinsèque d'activation de l'apoptose. DD : domaine de mort (Death Domain), DED : domaine effecteur de mort (Death Effector Domain), DISC : complexe de signalisation induisant la mort (Death Inducing Signalling Complex). 1 : reconnaissance d'un signal extracellulaire d'induction de l'apoptose par des récepteurs transmembranaires et trimérisation activatrice des récepteurs. 2 : recrutement de protéines adaptatrices par interaction entre les domaines de mort. 3 : recrutement de pro-caspases par interaction entre les domaines effecteurs de mort. 4 : activation de la caspase 8. 5 : cascade d'activation des caspases (Figure 12). 6 : induction de l'apoptose.

La voie intrinsèque de l'apoptose (Figure 14) peut être déclenchée par différents stress intracellulaires, notamment par une accumulation de lésions à l'ADN. Ces signaux induisent la formation de pores protéiques au niveau de la membrane mitochondriale externe par lesquels des facteurs pro-apoptotiques sont libérés dans le cytoplasme. Ces pores sont généralement formés par des protéines pro-apoptotiques de la famille protéique Bcl-2, tels que Bax et Bak. Constitutivement exprimées, ces protéines sont normalement inhibées par les membres anti-apoptotiques de la famille Bcl-2, tels que Bcl-2, Bcl-xL et Mcl-1. Les signaux de stress pro-apoptotiques permettent la libération et la multimérisation de Bax et Bak pour former des pores mitochondriaux. Ces pores permettent le relargage de plusieurs facteurs qui coopèrent pour déclencher l'apoptose. Le cytochrome c interagit avec la protéine Apaf1 dont il induit la transconformation. Cette protéine fixée au cytochrome c forme alors un homoheptamère, l'apoptosome qui active la pro-caspase 9 et déclenche la cascade d'activation des caspases. Parallèlement, d'autres facteurs comme SMAC/Diablo et Omi/HtraA2 inhibent des inhibiteurs de l'apoptose (IAP) et activent donc indirectement les caspases. Enfin, la

protéine AIF (Facteur Induisant l'Apoptose) libérée par la mitochondrie est transportée dans le noyau où elle favorise le clivage et la compaction de l'ADN (Adrain and Martin, 2001; Alberts et al., 2008; Cory and Adams, 2002).

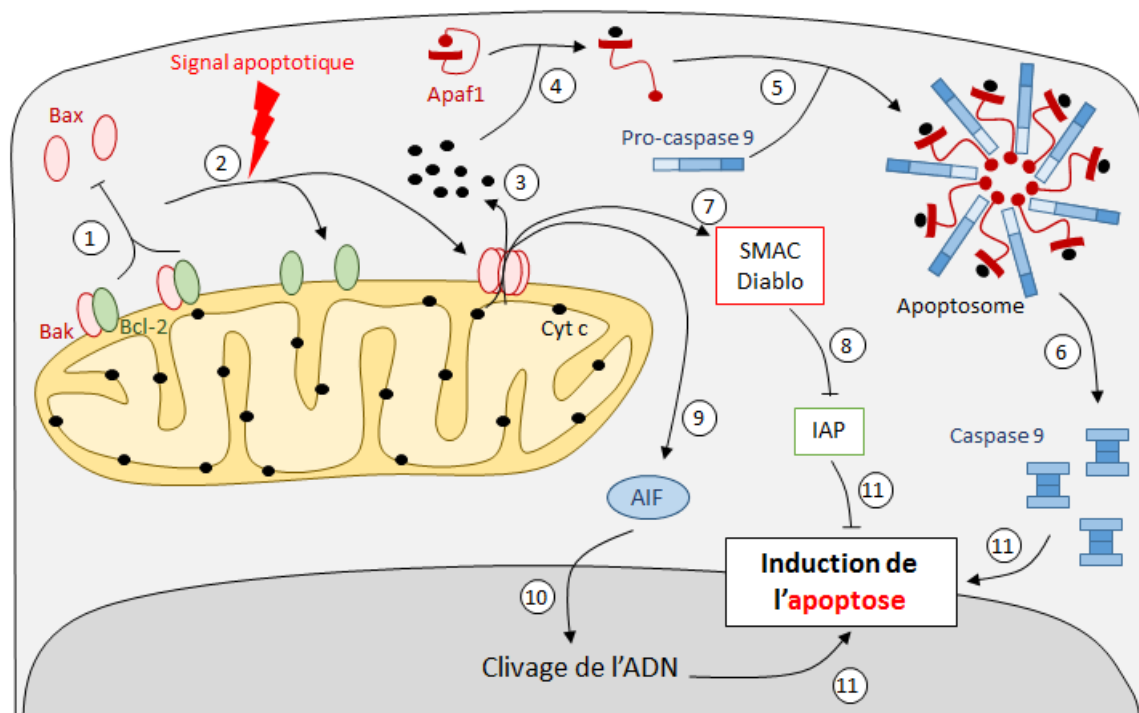


Figure 14 : représentation schématique de la voie intrinsèque d'activation de l'apoptose. 1 : les protéines Bak et Bax sont inhibées par les protéines anti-apoptotiques de la famille Bcl-2. 2 : un signal apoptotique intracellulaire lève l'inhibition de Bax et Bak. Ces deux protéines forment alors des pores mitochondriaux. 3 : du cytochrome c est relargué dans le cytoplasme depuis la mitochondrie. 4 : le cytochrome c induit la transconformation de la protéine Apaf1. 5 : sept molécules d'Apaf1 liées au cytochrome c s'associent avec sept pro-caspases 9 pour former l'apoptosome. 6 : la caspase 9 est activée. 7 : le facteur SMAC/Diablo est libéré hors de la mitochondrie. 8 : ce facteur inhibe les inhibiteurs de l'apoptose (IAP). 9 : la protéine AIF est libérée hors de la mitochondrie. 10 : AIF entre dans le noyau et favorise le clivage de l'ADN cellulaire. 11 : l'ensemble des processus de la voie intrinsèque concourt à l'induction de l'apoptose.

Au-delà de leur convergence vers l'activation des caspases, la voie extrinsèque et la voie intrinsèque de l'apoptose sont connectées via la protéine Bid. En effet, Bid est activée par la caspase 8 lors de la voie extrinsèque et active les protéines Bax et Bak de la voie intrinsèque. Elle favorise le relargage de facteurs pro-apoptotiques depuis la mitochondrie et donc la voie intrinsèque de l'apoptose (Korsmeyer et al., 2000).

Dans certains cancers, la résistance à l'apoptose est due à la surexpression de facteurs anti-apoptotique comme Bcl-2 (Leek et al., 1994; Lu et al., 1993). Mais dans la majorité des cas, des mutations sont retrouvées dans les gènes codant des facteurs clés de l'apoptose.

Dans environ 50% des cancers humains, la résistance à l'apoptose est due à une mutation dans le gène p53 (Alberts et al., 2008). En effet, ce gène code pour un facteur de transcription impliqué dans la régulation de différents processus cellulaires, dont l'apoptose. Il a été montré que p53 induit la transcription des facteurs pro-apoptotiques de la famille Bcl-2 et d'intermédiaires des voies intrinsèque et extrinsèque de l'apoptose mais également que p53 inhibe de façon directe ou non les inhibiteurs de l'apoptose. Ainsi, p53 est un activateur global de l'apoptose, agissant essentiellement mais pas uniquement sur la voie intrinsèque d'activation (Yu and Zhang, 2005).

Cette protéine est présente à des niveaux très faibles dans les cellules saines car elle est dégradée en parallèle de sa synthèse. Elle est stabilisée lorsque la cellule est confrontée à différents stress tels que les dommages à l'ADN, l'hypoxie, un manque de nucléotide, le stress oxydatif ou encore l'expression d'oncogènes. Cette augmentation du niveau de p53 cellulaire induit différentes réponses au stress, dont l'apoptose (Chipuk and Green, 2006). L'inactivation du gène correspondant dans des cellules transformées constitue donc une avancée drastique vers le phénotype cancéreux et p53 est considéré comme un suppresseur de tumeur majeur.

2/ Perte de contrôle du cycle cellulaire et augmentation de la prolifération

La prolifération d'une cellule saine est finement régulée par différents signaux extracellulaires qui permettent de contrôler le nombre de cellules au sein de chaque tissu. Les cellules cancéreuses sont à la fois affranchies de ces mécanismes de régulation et dans un état d'activation permanent de leur prolifération cellulaire. Autrement dit, indépendamment des signaux extérieurs, les processus de croissance cellulaire, réplication du génome et division cellulaire sont globalement activés dans les cellules cancéreuses, tout en restant coordonnés entre eux.

Des mutations dans des dizaines de gènes peuvent être à l'origine de la non sensibilité des cellules cancéreuses aux signaux extérieurs de contrôle de la prolifération (Hanahan and Weinberg, 2000). En effet, ces signaux activent différentes voies de signalisation et chacune d'entre elle peut être interrompue par la suppression d'un de ses intermédiaires protéiques. Le détail des différentes voies de signalisation et des intermédiaires pour lesquels des mutations ont été observées ne sera pas repris ici.

Mais les cellules cancéreuses ne sont pas seulement affranchies des signaux de régulation externes. Elles sont également dans un état d'activation permanent de la prolifération cellulaire. Ainsi les mutations qui ciblent directement les étapes clés de contrôle du cycle cellulaire sont retrouvées dans de nombreux cancers.

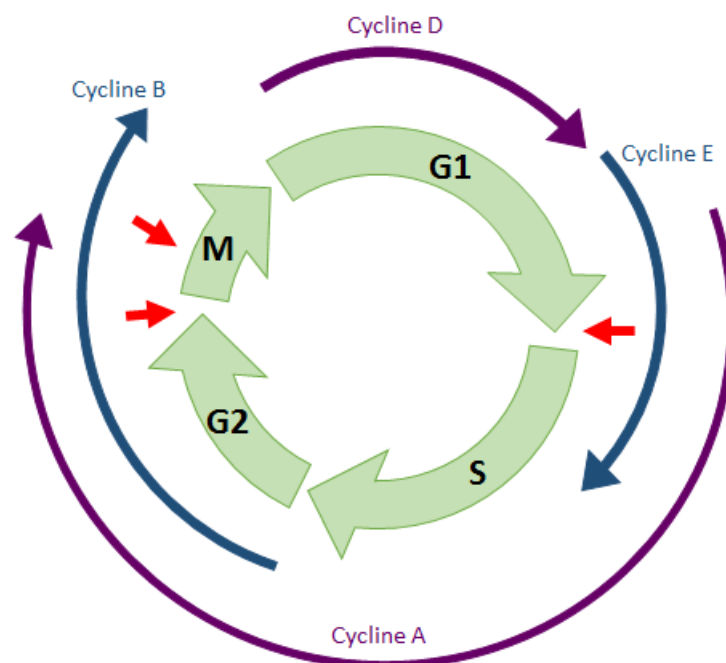


Figure 15 : schéma des cyclines exprimées pendant les différentes phases du cycle cellulaire chez les vertébrés. Les trois flèches rouges indiquent les points de contrôle.

Le cycle cellulaire repose sur une succession de cycles de division comprenant chacun quatre phases : la phase G1 (croissance cellulaire), la phase S (réplication de l'ADN), la phase G2 (croissance cellulaire et préparation de la mitose) et la phase M (mitose). La régulation de la progression du cycle cellulaire implique de nombreux mécanismes et est articulée autour de trois points de contrôle (ou checkpoints) temporels : la transition G1/S, la transition G2/M et un point de contrôle au sein de la mitose, entre les étapes de métaphase et d'anaphase, c'est-à-dire avant la séparation des chromosomes des deux cellules filles (Figure 15) (Molinari, 2000). Chaque point de contrôle ne peut être passé que si la cellule détecte que les étapes préalables du cycle cellulaire ont eu lieu correctement. Par exemple, la transition G2/M est bloquée si le génome n'a pas été entièrement dupliqué.

A ces trois points de contrôle peut être ajoutée l'étape d'entrée dans le cycle cellulaire. En effet, la plupart des cellules de notre organisme ne se divisent pas et sont dans une phase dite G0. L'enclenchement du cycle cellulaire et donc un préalable à leur division cellulaire (Johnson and Walker, 1999).

Bien que chaque point de contrôle du cycle cellulaire ne puisse être passé que si les étapes précédentes se sont bien déroulées, les protéines régulant la progression du cycle cellulaire sont différentes de celles qui sont activement impliquées dans ses quatre phases. En effet, la régulation du cycle cellulaire est médiée par des protéines kinases dépendantes des cyclines (CDK). Comme leur nom l'indique, ces kinases ont la particularité de n'être actives que lorsqu'elles sont associées à des protéines appelées cyclines (Morgan, 1995). Or, si le niveau d'expression des CDK est constant, celui des cyclines varie au cours du cycle cellulaire. Ainsi, chez les vertébrés, il existe quatre types de cyclines qui sont exprimées uniquement lors de certaines phases du cycle cellulaire : les cyclines D et E pendant la phase G1, la cycline A de la phase S au début de la mitose et la cycline B, de la phase G2 à la phase M (Figure 15). Chacune de ces cyclines n'interagissant qu'avec certaines CDK, chaque étape du cycle cellulaire est caractérisée par la présence de dimères cycline/CDK spécifiques (Tableau 3) (Johnson and Walker, 1999).

Cyclines	CDK associées
D	4, 6
E	2
A	2, 1
B	1

Tableau 3 : rappel des complexes cycline/CDK formés dans les cellules de vertébrés

A cette régulation temporelle par l'expression de différentes cyclines s'ajoutent des régulations directes par des protéines capables d'inhiber ou d'activer les CDK. Ces régulateurs sont généralement exprimés dans des conditions particulières (dommages à l'ADN, progression au cours de la mitose...) et permettent le contrôle de la progression du cycle cellulaire en fonction des conditions réelles auxquelles est confrontée la cellule (Alberts et al., 2008; Morgan, 1995).

A chaque point de contrôle du cycle cellulaire est associée une série de réactions permettant d'enclencher la phase suivante du cycle. Ces réactions aboutissent à la régulation de l'expression de gènes spécifiques des différentes phases du cycle cellulaire et sont finement régulées. Or un certain nombre de mutations entraînant le développement de cancers perturbent justement ces réactions de transition.

C'est notamment le cas des mutations inactivant le suppresseur de tumeur *Rb-1*. La protéine Rb est en effet impliquée dans le blocage de l'entrée en phase S : elle inhibe les facteurs de la famille E2F qui sont nécessaires à la transcription des gènes permettant l'entrée en phase S (Figure 16, gauche). Rb est elle-même inhibée par le complexe cycline D/CDK4, spécifiquement exprimé au

moment de la transition G1/S. L'inhibition exercée par Rb est donc normalement levée uniquement dans les cellules qui sont entrées dans le cycle cellulaire et précisément au moment de l'entrée en phase S (Figure 16, centre). La mutation de *Rb-1* dans les cellules cancéreuses permet à ces dernières d'activer la phase S indépendamment de tout mécanisme de contrôle (Chinnam and Goodrich, 2011; Dick and Rubin, 2013). Cela a pour double conséquence d'augmenter leur taux de prolifération et de favoriser leur instabilité génétique. En effet, la réplication peut alors avoir lieu dans des conditions inadaptées induisant l'augmentation du taux d'erreurs lors de la réplication du génome.

Le mécanisme de contrôle impliquant les protéines Rb, cycline D et CDK4 est extrêmement important dans la cellule. En effet, le complexe cycline D/CDK4 peut être inhibé par la protéine p16, qui est produite en cas de stress cellulaire (Figure 16, droite). Ainsi, en cas de stress et d'expression de p16, c'est le blocage exercé par Rb qui empêche la progression du cycle cellulaire vers la phase S (Li et al., 2011; Rayess et al., 2012). Les cellules cancéreuses exprimant une protéine Rb fonctionnelle possèdent généralement d'autres mutations altérant le fonctionnement de ce mécanisme de contrôle. Certaines surexpriment la cycline D ou la CDK4 (Dickson, 2014), d'autres possèdent des mutations ou des modifications épigénétiques inactivant le gène *p16* (Li et al., 2011).

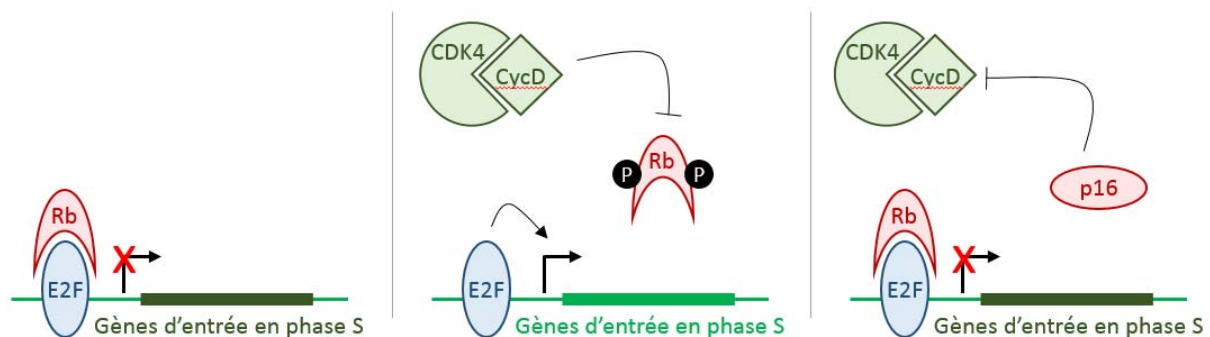


Figure 16 : contrôle du cycle cellulaire par la protéine Rb. Gauche : Rb inhibe le facteur de transcription E2F et empêche l'expression des gènes d'entrée en phase S. Centre : le complexe cycline D/CDK4 exprimé en phase G1 inhibe Rb par phosphorylation et permet l'expression des gènes d'entrée en phase S. Droite : en cas de stress cellulaire, la protéine p16 inhibe le complexe cycline D/CDK4 et empêche la progression du cycle cellulaire.

Enfin, le suppresseur de tumeur *p53* est également impliqué dans l'inhibition de l'entrée en phase S du cycle cellulaire. En effet, il induit la transcription de p21, un inhibiteur de CDK bloquant la levée de l'inhibition exercée par Rb. Ce mécanisme est activé en cas de stress cellulaire et donc d'augmentation de la quantité de p53 (Giono and Manfredi, 2006). Il empêche l'entrée de la cellule en phase S si les conditions ne sont pas propices au bon déroulement de cette dernière. Les mutations inactivant *p53* favorisent donc également la progression vers la phase S du cycle cellulaire, dans des conditions susceptibles d'augmenter l'instabilité génétique des cellules cancéreuses.

3/ Altération de la réparation des dommages à l'ADN

Si l'instabilité génétique des cellules cancéreuses est en partie due à leur progression forcée vers la phase S du cycle cellulaire, et donc à l'accumulation d'erreurs au cours de la réplication de leurs génomes, elle serait moindre si ces cellules possédaient des mécanismes de réparation efficace des dommages à l'ADN. Outre les oncogènes et les suppresseurs de tumeurs, le dysfonctionnement des gènes de maintenance du génome joue un rôle clé dans le développement du cancer.

Il existe un certain nombre de mécanismes de réparation des dommages à l'ADN dans les cellules humaines, qui sont adaptés à la réparation des différents types de lésions susceptibles de se produire. Dans la plupart des cas les mécanismes de réparation des dommages à l'ADN sont redondants et une même lésion pourrait être prise en charge par plusieurs voies de réparation. Cette

redondance fonctionnelle représente un niveau de sécurité supplémentaire pour la cellule et diminue la probabilité de stabilisation d'une mutation.

Les mutations ponctuelles induisant des mésappariements peuvent ainsi être réparées par excision de base, c'est-à-dire clivage de la base azotée anormale puis clivage du désoxyribose mono-phosphate restant (Kim and Wilson III, 2011), ou par excision de nucléotides, c'est-à-dire clivage d'une série de nucléotides comprenant le nucléotide muté (Schärer, 2013) (Figure 17). Ces deux mécanismes génèrent des coupures simple brin de l'ADN. Le remplacement du ou des nucléotides manquant est alors réalisé par une ADN polymérase à partir du brin complémentaire. Les fragments néosynthétisés sont ensuite incorporés au sein du brin endommagé par l'action d'une ADN ligase (Kim and Wilson III, 2011; Schärer, 2013).

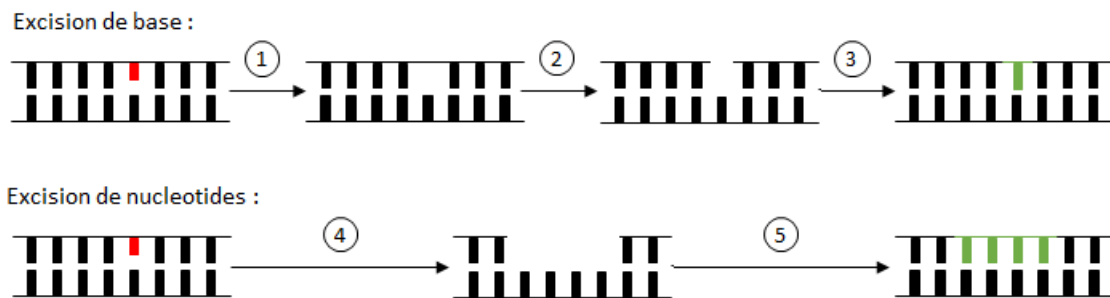


Figure 17 : principe de la réparation des mésappariements par excision de base ou de nucléotides. Les rectangles noirs représentent les nucléotides originels, les rouges les nucléotides mutés, les verts les nucléotides néosynthétisés. 1 : clivage de la base azotée modifiée. 2 : clivage du désoxyribose mono-phosphate correspondant. 3 : synthèse du nucléotide manquant par complémentarité et insertion dans l'ADN. 4 : excision d'une série de nucléotides contenant le muté. 5 : synthèse par complémentarité et insertion dans l'ADN.

Ces deux mécanismes sont relativement rapides mais ne permettent de réparer que les mutations ponctuelles concernant un nombre restreint de nucléotides. Le dommage à l'ADN susceptible de poser le plus de problèmes à la cellule est la rupture double brin de l'ADN. Celle-ci peut être causée par des radiations ionisantes, par des erreurs de réplication et par différents métabolites cellulaires dont les agents oxydants. Sa réparation est plus complexe car aucun modèle permettant une synthèse par complémentarité n'est disponible.

Dans la majorité des cas et notamment dans les cellules somatiques, ce type de lésion est réparé par jonction d'extrémités non homologues (NHEJ) (Figure 18). Les deux sites de clivages sont reconnus par des hétérodimères des protéines Ku70 et Ku80. Différents autres facteurs sont recrutés aux sites de clivages et, si nécessaire, les extrémités libres de l'ADN sont clivées jusqu'à obtention d'une microhomologie entre les deux extrémités libres qui permet leur ligation (Burma et al., 2006; Rassool, 2003). Ce mécanisme de réparation est rapide mais conduit à la délétion d'un fragment d'ADN. L'apparition de ces délétions est plus ou moins dommageable selon l'organisme (proportion codante du génome) et le type cellulaire (cellule somatique ou germinale). Il est de toute façon le principal mécanisme exploitable par les cellules pour réparer les ruptures double brin de l'ADN en dehors des phases S et G2 (Alberts et al., 2008).

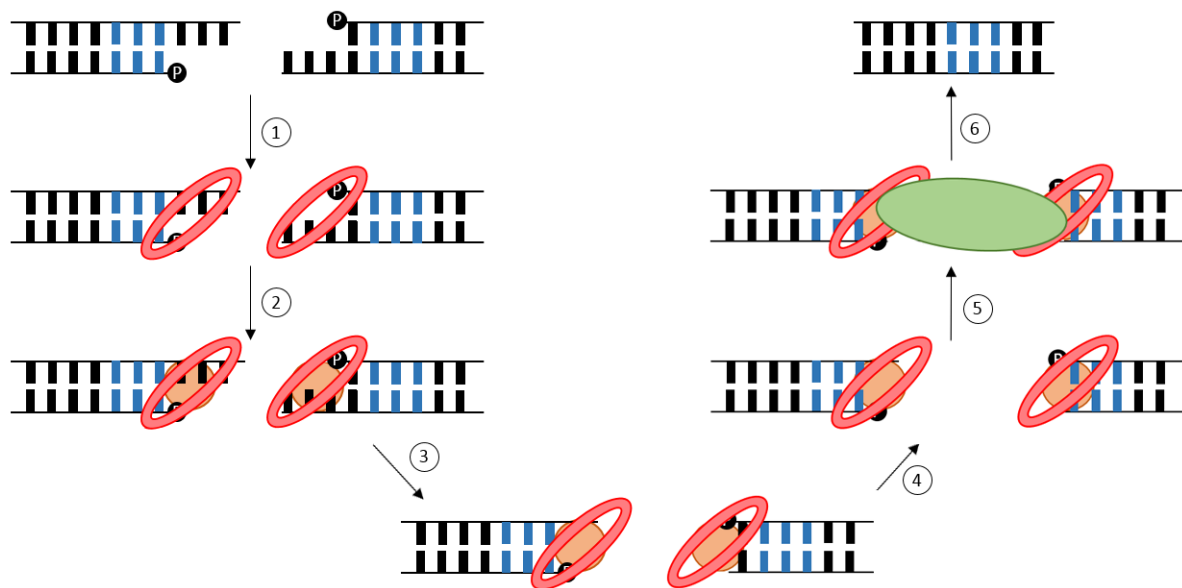


Figure 18 : principe de la réparation des ruptures double brin par NHEJ. Les nucléotides en bleus correspondent à une micro-homologie de séquence. 1 : reconnaissance du clivage par Ku70/Ku80. 2 : recrutement de la machinerie de réparation. 3 : clivage des extrémités non appariées. 4 : en l'absence d'homologie, poursuite du clivage. 5 : recrutement de la machinerie de ligation. 6 : ligation avec délétion par rapport à la séquence d'origine.

En effet, entre la réplication du génome cellulaire et la mitose, l'information génétique est présente en double dans la cellule sous la forme de deux chromatides sœurs. Dans ce contexte, les ruptures double brin de l'ADN d'une des chromatides sœurs peuvent être réparées en utilisant la seconde chromatide comme modèle via le mécanisme de recombinaison homologue (Figure 19). Dans un premier temps, les deux extrémités du site de clivage sont dégradées dans le sens 5'-3'. Un des deux brins d'ADN clivé va alors entrer en contact avec la double hélice intacte de la chromatide sœur. En présence d'homologie de séquence entre l'ADN des deux chromatides, les deux brins de la double hélice intacte se séparent localement et l'un d'entre eux s'apparie avec le brin clivé. La partie manquante de ce brin d'ADN est alors synthétisée par complémentarité avec celui de la chromatide sœur. Les brins d'ADN des deux chromatides se séparent ensuite, la chromatide intacte retrouvant son appariement originel. La partie néosynthétisée de la chromatide clivée peut alors s'apparier avec la séquence complémentaire du second brin clivé, qui pourra lui-même être complété par complémentarité avec le premier brin. Une ADN ligase intervient après la seconde étape de synthèse pour permettre la ligation de l'ADN néosynthétisé avec le brin correspondant (San Filippo et al., 2008). Ce mécanisme de réparation est plus complexe et ne peut avoir lieu que pendant certaines phases du cycle cellulaire, mais il ne conduit généralement pas à l'apparition de mutation. En revanche, si, en raison de l'organisation spatiale de l'ADN dans le noyau, une chromatide du second chromosome de la paire est utilisée comme modèle au lieu de la chromatide sœur, la recombinaison homologue peut conduire à une perte d'hétérozygotie (Carr and Lambert, 2013).

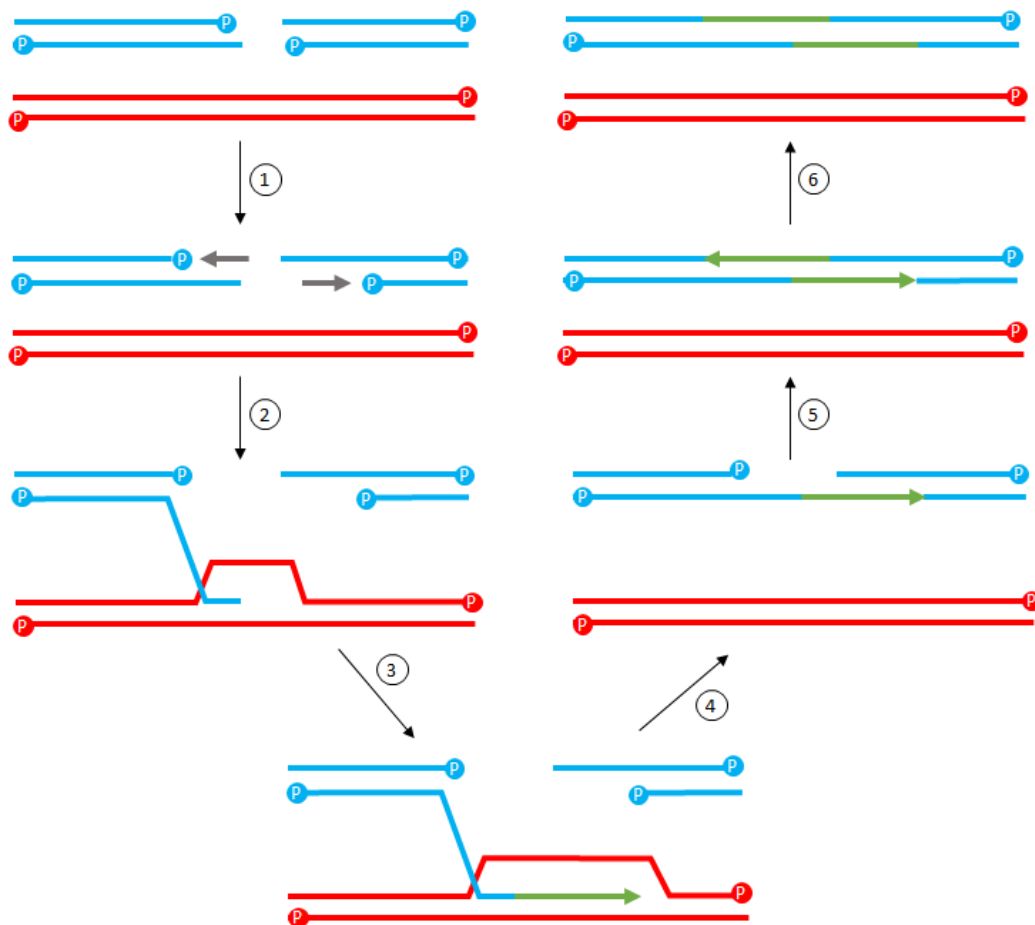


Figure 19 : principe de la réparation des ruptures double brin par recombinaison homologue. Les deux brins de chaque chromatide sœur sont représentés de la même couleur (bleu pour la double hélice rompue, rouge pour l'autre). L'ADN néosynthétisé est représenté en vert. 1 : digestion 5'-3' à partir de la rupture double brin. 2 : appariement d'un des brins rompus avec un brin de la chromatide intacte (invasion de brin). 3 : synthèse par complémentarité. 4 : séparation des deux chromatides et appariement du deuxième brin rompu avec l'ADN néosynthétisé. 5 : synthèse par complémentarité. 6 : ligation de l'ADN néosynthétisé.

Un certain nombre de facteurs sont impliqués dans l'ensemble de ces mécanismes de réparation et la liste exhaustive n'en sera pas dressée ici. En revanche il est intéressant de noter que la mutation de certains de ces facteurs peut favoriser le développement de cancers. Ainsi, BRCA-1 et BRCA-2 sont impliqués dans les premières étapes de la recombinaison homologue et leur inactivation est à l'origine de certains cancers du sein (Prakash et al., 2015).

De plus la surexpression de certains oncogènes peut altérer le fonctionnement de ces mécanismes de réparation. C'est notamment le cas de *myc*. Ce facteur de transcription est un régulateur bien caractérisé de la prolifération cellulaire et de l'apoptose. Il a récemment été montré que sa surexpression, entre autres dans les cellules cancéreuses, ralentit la réparation des ruptures double brins de l'ADN par NHEJ et par recombinaison homologue (Ambrosio et al., 2015; Li et al., 2012). La surexpression de *myc* augmente par conséquent le nombre de ces ruptures double brin et l'instabilité génétique des cellules cancéreuses.

Les connaissances disponibles concernant le développement des cancers progressent régulièrement et de nombreuses études sont actuellement menées pour mieux comprendre les mécanismes de l'oncogenèse. Si certains processus, comme la formation des métastases, restent encore mal compris, l'ensemble des données disponibles tend à montrer que la réussite thérapeutique

dépend non pas de la compréhension du cancer en général, mais d'un cancer en particulier, celui dont souffre le patient (Heuckmann and Thomas, 2015).

En effet, de nombreux profils mutationnels sont susceptibles de conduire au développement d'un cancer et cette convergence phénotypique masque une grande diversité dans les mécanismes mis en place. Le traitement des cancers est encore essentiellement basé sur le ciblage de caractéristiques générales des cellules cancéreuses, notamment leur rapidité de prolifération. Ce modèle thérapeutique évolue progressivement vers des traitements personnalisés, adaptés au profil mutationnel de chaque cancer. Ainsi, seules les patientes souffrant de tumeurs mammaires exprimant les récepteurs aux œstrogènes et à la progestérone sont orientées vers un traitement par hormonothérapie. Sur ce principe, des approches de plus en plus complexes permettant d'affiner les traitements sont en cours de conception (Le Du et al., 2015). Si ce type démarche est évidemment prometteur en terme de pronostic pour les patients, il nécessite cependant une compréhension des mécanismes moléculaires à l'origine de l'oncogenèse plus détaillée et exhaustive que l'actuelle.

III/ SAMHD1, histoire et contexte d'une découverte majeure

A/ Le VIH : structure virale et adaptation à l'environnement cellulaire

1/ Les VIH, des lentivirus à l'origine d'une pandémie

a/ Le SIDA : de la découverte du VIH à la gestion d'une pandémie

Les premiers patients atteints par le Syndrome d'ImmunoDéficiency Acquis (SIDA) ont été identifiés en 1981 par le CDC (Center for Disease Control) d'Atlanta comme atteints par une pneumonie et une déplétion des lymphocytes T. Le SIDA a été défini comme un ensemble de symptômes causés par la perte de fonctionnalité du système immunitaire, ce qui favorise le développement d'infections opportunistes et les rend mortelles. Des symptômes similaires ont été observés chez des patients français (Barré-sinoussi, 2003) et il a fallu moins de deux ans à des chercheurs français pour identifier un nouveau virus présent chez les malades (Barre-Sinoussi, F., J. C. Chermann, 1983). Ce virus, proche des HTLV-1 et HTLV-2 précédemment caractérisés, a été classé dans la catégorie des rétrovirus en raison de sa morphologie et de son association avec une activité transcriptase inverse (Barre-Sinoussi, F., J. C. Chermann, 1983). Le séquençage de son génome a confirmé son appartenance à cette famille virale, le plaçant dans le genre plus restreint des lentivirus, tout en soulignant sa différence par rapport aux virus déjà connus (Wain-Hobson et al., 1985). D'autres équipes avaient isolé des virus à partir de prélèvements de patients atteints du SIDA et il a finalement été montré que ces différents isolats n'étaient que des souches différentes du même virus (Ratner et al., 1985) : le Virus de l'Immunodéficience Humaine (VIH).

Bien qu'isolé et identifié comme agent causal du SIDA il y a une trentaine d'années, beaucoup de choses restent à comprendre concernant ce virus et la mise au point d'un vaccin est encore un défi à relever. De fait, le VIH est à l'origine d'une pandémie et si les traitements permettent désormais d'améliorer significativement la qualité de vie des malades, il reste impossible de les guérir. D'après le dernier rapport du programme des Nations Unies concernant le SIDA, il y avait 35 millions de personnes infectées par le VIH en 2013 et même si le nombre de contaminations annuelles a diminué de 38% depuis 2001, 2,1 millions de personnes ont été contaminées en 2013 (UNAIDS, 2014). La disponibilité des traitements reste très variable selon les zones géographiques et aujourd'hui seuls 37% des patients sont traités. Et même pour les patients sous traitement, le SIDA demeure une maladie mortelle. Si la mortalité annuelle a chuté de 35% depuis 2005, 1,5 million de personnes sont décédées en raison du SIDA en 2013 (UNAIDS, 2014). Depuis le début de l'épidémie, le SIDA a touché 78 millions de personnes et entraîné 39 millions de décès, ce qui en fait un des plus redoutables agents pathogènes de l'histoire de l'humanité.

A l'heure actuelle, le meilleur moyen de lutter contre le SIDA reste la prévention mais les progrès réalisés en termes de thérapies anti-rétrovirales méritent d'être soulignés. Ces thérapies permettent désormais d'empêcher les personnes infectées de transmettre le virus et retardent voire bloquent l'apparition des symptômes (Cohen, 2012). Le principal objectif actuel est d'étendre et d'améliorer leur distribution dans les pays en voie de développement, où les malades sont plus nombreux que dans les pays développés mais manquent de moyens financiers pour se soigner (Ebrahim and Mazanderani, 2013). Cela n'empêche pas en parallèle de chercher à améliorer les thérapies actuellement disponibles et les recommandations de l'Organisation Mondiale de la Santé sont régulièrement mises à jour en fonction des derniers résultats cliniques (Olubajo et al., 2014).

b/ Organisation des particules et des génomes viraux

Le VIH appartient à la famille des rétrovirus, qui rassemble des virus enveloppés d'une centaine de nanomètres de diamètre dont le génome se présente sous la forme de deux copies d'ARN simple brin. Cet ARN est de polarité positive, c'est-à-dire que sa séquence est identique (et non

complémentaire) à celle utilisée comme support lors de la traduction des protéines virales. Il est rétrotranscrit en ADN viral au cours du cycle de réplication des rétrovirus par une transcriptase inverse virale. Parmi les rétrovirus, les lentivirus ont, comme leur nom l'indique, la particularité d'entraîner des maladies avec une longue période d'incubation. Contrairement aux autres rétrovirus, leur cycle de réplication est indépendant du démantèlement de l'enveloppe nucléaire au cours de la mitose.

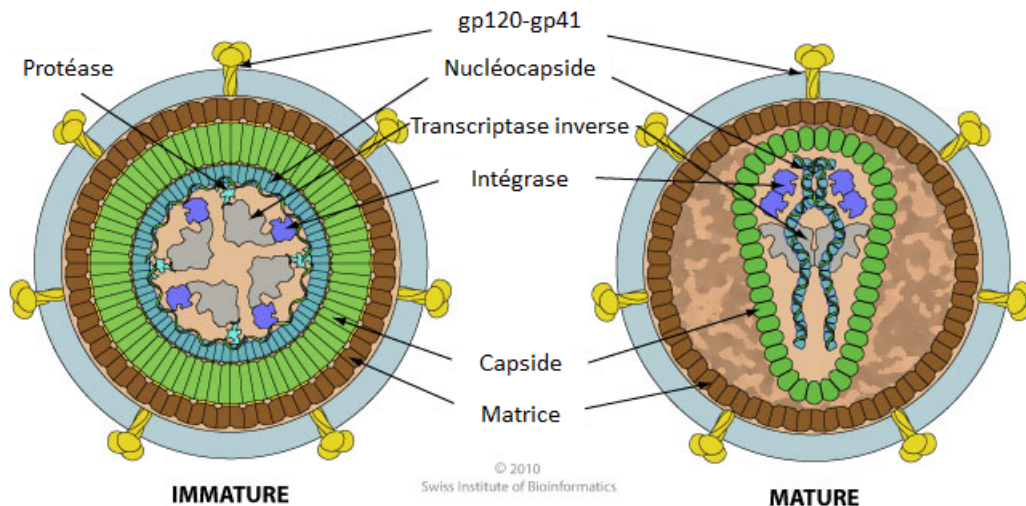


Figure 20 : schéma de l'organisation des particules rétrovirales. Source: ViralZone www.expasy.org/viralzone, Swiss Institute of Bioinformatics.

Les particules virales des rétrovirus ont une organisation conservée (Figure 20) qui n'est acquise que suite à une étape de maturation catalysée par la protéase virale après le bourgeonnement des virions néo synthétisés. La particule virale est délimitée par une membrane lipidique provenant de la membrane plasmique de la cellule productrice, au niveau de laquelle sont localisées les protéines virales d'enveloppe : la gp120 (surfactive) et la gp41 (transmembranaire). La protéine de matrice (MA) virale forme une structure sous-jacente à cette enveloppe lipidique dans laquelle elle s'ancre par myristoylation. Au sein de la particule, la protéine de capside (CA) virale forme une seconde structure de forme trapézoïdale caractéristique. Celle-ci contient les deux brins d'ARN viraux de polarité positive empaquetés dans des polymères de protéine de nucléocapside (NC), des enzymes virales, notamment l'intégrase (IN), la transcriptase inverse (RT) et la protéase (PR) mais aussi certaines protéines virales dites accessoires, des protéines cellulaires et des composés nécessaires à la transcription inverse du génome viral (Scarlata and Carter, 2003).

Le génome du VIH, pris ici comme modèle de génome de lentivirus, possède 9 cadres ouverts de lectures répartis sur un peu moins de 10kb d'acide nucléique. Il code pour trois grands types de protéines : les protéines de structure, les protéines régulatrices et les protéines accessoires (Figure 21) (Scarlata and Carter, 2003).

Les protéines de structure sont codées par trois cadres ouverts de lecture. Gag (Group Antigen Specific) est transcrit sous la forme d'un ARNm non épissé, lui-même traduit sous la forme d'un précurseur conséquemment clivé par la protéase virale pour former les protéines structurales du virus (MA, CA et NC précédemment évoquées ainsi que le peptide p6, présent uniquement chez les lentivirus de primates). Pol (polymérase) est traduit sous la forme d'une polyprotéine Gag-Pol via un décalage des ribosomes sur un ARNm non épissé. Il code pour les enzymes virales impliquées dans la transcription inverse du génome viral (RT), son intégration dans le génome cellulaire (IN) et la maturation des particules virales (PR). Enfin, Env (enveloppe) est transcrit sous la forme d'un ARNm mono-épissé lui-même traduit sous la forme d'un précurseur gp160 qui est clivé en gp120 et gp41 par une protéase cellulaire.

Les protéines régulatrices du VIH Tat et Rev sont transcrites sous forme d'ARNm épissés deux fois, ce qui permet leur export nucléaire par la même voie que les ARNm cellulaires. Rev (regulator of virion protein expression) est impliqué dans l'export des autres ARNm viraux (cf I/D/1) tandis que Tat (Trans-Activator of Transcription) permet d'induire la transcription du génome viral.

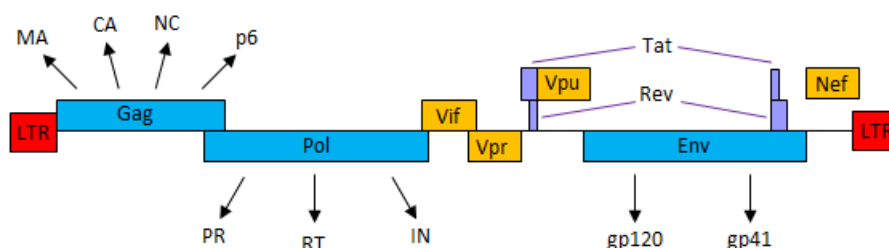


Figure 21 : représentation schématique du génome du VIH-1. Les gènes codant les protéines de structure sont indiqués en bleu, ceux codant pour les protéines accessoires en jaune et ceux codant pour les protéines régulatrices en violet.

Ce génome code également pour un certain nombre de protéines accessoires traduites à partir d'ARNm mono ou bi-épissés. Elles ont été qualifiées d'accessoires car elles ne sont pas nécessaires à la réplication virale dans des cellules permissives *in vitro* mais elles jouent en réalité un rôle fondamental dans la pathogenèse *in vivo*. Si les autres protéines décrites ci-dessus sont conservées chez tous les lentivirus, les protéines accessoires diffèrent selon le lentivirus considéré. Elles constituent un des marqueurs permettant de décrire l'histoire évolutive des souches virales et notamment de comprendre les origines de la transmission à l'homme de virus d'immunodéficience initialement simiens.

Catégorie	Nom	Abréviation	Précurseur	Fonction proposée
Protéines de structure	Capside	CA	Gag	Structure
	Matrice	MA	Gag	Structure
	Nucléocapside	NC	Gag	Structure (protection de l'ARN viral)
Protéines d'enveloppe	gp120		Env	Reconnaissance du récepteur et du co-récepteur cellulaires
	gp41		Env	Fusion des membranes virale et plasmique
Protéines enzymatiques	Intégrase	IN	Pol	Intégration du génome viral dans le génome cellulaire
	Protéase	PR	Pol	Maturation des particules virales
	Transcriptase inverse	RT	Pol	Transcription inverse du génome viral
Protéines régulatrices	Tat			Activation de la transcription du génome viral
	Rev			Export des ARNm viraux non ou mono épissés
Protéines accessoires	Nef			Augmentation de l'infection des cellules non permissives (cible BST-2)
	Vif			Augmentation de l'infection des cellules non permissives (cible APOBEC3G)
	Vpr			Augmentation de l'infection des cellules non permissives
	Vpu			Augmentation de l'infection des cellules non permissives (cible BST-2)
	Vpx			Augmentation de l'infection des cellules non permissives (cible SAMHD1)

Tableau 4 : tableau récapitulant les différentes protéines du VIH. Seules les fonctions principales et évoquées dans ce manuscrit y ont été rappelées.

c/ Classification et origines des VIH

L'emploi du singulier pour parler du VIH relève en réalité de la simplification. En effet, sans même considérer la variabilité des génomes viraux chez un seul patient infecté, il existe plusieurs types de VIH.

Après l'identification du premier VIH, dit VIH-1, chez les patients souffrant de SIDA en Europe et aux Etats-Unis (Barre-Sinoussi, F., J. C. Chermann, 1983), un second virus morphologiquement proche mais possédant un profil antigénique différent a été identifié comme agent causal du SIDA chez des patients d'Afrique de l'ouest et baptisé VIH-2 (Clavel et al., 1986). En plus de ces deux grands types de VIH, les analyses de séquences de différentes souches de VIH-1 ont mis en évidence l'existence de quatre groupes distincts. Le groupe M (main) est le plus répandu et est à l'origine de la pandémie actuelle de SIDA, les groupes N et P n'ont été identifiés que chez un nombre réduit de patients camerounais et le groupe O, bien qu'également circonscrit dans une zone géographique réduite, a causé plus de 100000 infections en Afrique de l'ouest. Des clades distincts ont été identifiés au sein du groupe M du VIH-1 et parmi les souches de VIH-2 (Hahn et al., 2000).

A ces différentes catégories de virus infectant l'homme, il faut ajouter un certain nombre de virus simiens, infectant essentiellement les singes africains. En effet, des virus morphologiquement proches du VIH ont été identifiés chez différentes espèces de singes. Une même souche virale peut d'ailleurs ne pas être pathogène pour l'espèce de singe avec laquelle elle a co-évolué et conduire au développement de SIDA si elle infecte une autre espèce. On regroupe ces virus sous le terme de Virus de l'Immunodéficience Simienne (Chahroudi et al., 2012).

L'analyse des séquences des génomes de l'ensemble de ces VIH et VIS permet de tracer des arbres phylogénétiques viraux. De façon intéressante, ceux-ci ont montré que le VIH-2 et les quatre groupes de VIH-1 ne forment pas un groupe monophylétique mais sont apparentés à des VIS différents. Cela montre que l'existence de plusieurs virus humains différents n'est pas due à l'évolution divergente d'une même souche ancestrale pathogène pour l'homme mais à la combinaison de différents passages de la barrière d'espèce ayant permis la contamination de l'homme par plusieurs virus simiens (Hahn et al., 2000).

Des analyses phylogénétiques combinées aux données géographiques de localisation des différentes souches virales humaines et simiennes ont ainsi permis de montrer que le VIH-2 résulte de la transmission à l'homme du VISmm, virus infectant de façon non pathogène les mangabeyes enfumés mais conduisant au SIDA les macaques infectés. Les groupes M et N du VIH-1 résultent quant à eux de deux transmissions à l'homme du VIScpz, un VIS pathogène chez le chimpanzé issu de la combinaison chez ce singe d'un virus non pathogène chez les cercopithèques et d'un virus non pathogène chez les mangabeyes couronnés. Le groupe P du VIH-1 résulte de la transmission à l'homme du VISgor, un VIScpz ayant évolué pour infecter les gorilles (Sharp and Hahn, 2011). Enfin, il a récemment été montré que le groupe O du VIH-1 a également pour origine une contamination de l'homme par le gorille (D'arc et al., 2015).

En plus des VIS susmentionnés, il existe un grand nombre de VIS qui infectent les populations de singes mais n'ont jamais été transmis à l'homme (Beer et al., 1999; Chahroudi et al., 2012; Sharp and Hahn, 2011).

Si on compare l'organisation globale des génomes de l'ensemble des VIH et VIS, trois schémas d'organisation, différents en termes de protéines accessoires, sont identifiables (Figure 22). L'ensemble des virus de l'immunodéficience codent pour les protéines Vif, Vpr et Nef. Le VIH-1 et les VIS apparentés codent en plus pour la protéine Vpu tandis que le VIH-2 et les VIS apparentés codent pour la protéine Vpx et qu'une partie des VIS ne codent que pour les trois protéines accessoires communes (Beer et al., 1999).

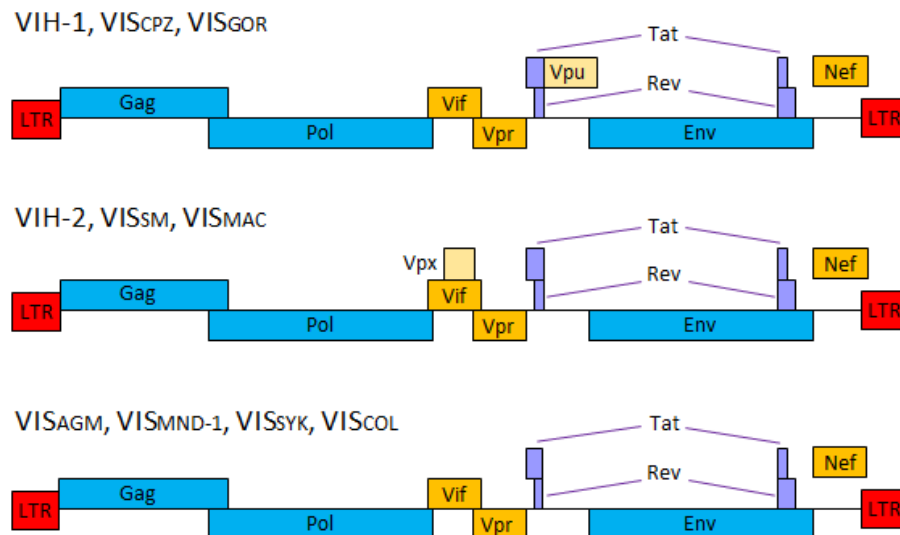


Figure 22 : représentation schématique des génomes des VIH et des virus simiens apparentés.

2/ Réplication dans les cellules hôtes

a/ Cycle réplicatif du VIH

Le cycle de réplication du VIH peut-être divisé en deux phases : une phase précoce, de l'entrée du virus dans la cellule cible à l'intégration du génome viral dans le génome cellulaire, et une phase tardive, de l'expression des gènes viraux à la production de particules virales (Figure 23).

L'entrée du VIH dans ses cellules cibles est médiée par l'interaction entre la protéine d'enveloppe gp120 et le récepteur cellulaire CD4. Celle-ci induit une transconformation de la gp120 qui la rend capable d'interagir avec un co-récepteur qui, selon les souches virales, peut être CCR5 ou CXCR4. La reconnaissance d'un de ces récepteurs aux chimiokines par la gp120 permet un changement de conformation de la gp41 conduisant à la fusion de la membrane virale avec la membrane plasmique (Klasse, 2012).

Suite à l'entrée de la capsid virale dans le cytoplasme, trois étapes clés du cycle viral ont lieu : la décapsidation, la transcription inverse, et le transport jusqu'aux pores nucléaires. L'orchestration de ces trois étapes dans le temps et dans l'espace reste encore aujourd'hui un sujet de controverse. Les données expérimentales tendent à soutenir l'hypothèse selon laquelle une décapsidation partielle est nécessaire pour permettre la transcription inverse mais que cette dernière ne peut cependant avoir lieu qu'au sein de la capsid pour limiter la reconnaissance du matériel génétique viral par les senseurs de l'immunité innée. Ainsi la décapsidation pourrait n'avoir lieu qu'au niveau des pores nucléaires pour permettre le passage du complexe de pré-intégration directement de la capsid au nucléoplasme. Mais ce modèle reste discuté et de nombreuses études soulignent l'importance de différentes protéines cellulaires dans les processus cytoplasmiques de la phase précoce du cycle viral (Arhel, 2010; Fassati, 2012).

Les étapes susmentionnées permettent la mise en place d'un Complexe de Pré-Intégration (PIC) contenant l'ADN double brin viral, des protéines virales et des protéines cellulaires. L'import

nucléaire de ce complexe est indépendant de la mitose car il se fait via les pores nucléaires et ne nécessite pas de disruption de l'enveloppe nucléaire. Néanmoins, même si les données expérimentales permettent d'identifier certains acteurs clés de ce processus, le mécanisme d'import nucléaire impliqué est encore mal compris (Matreyek and Engelman, 2013).

Une fois le PIC importé dans le noyau, l'intégrase virale permet l'insertion de l'ADN viral dans le génome cellulaire. Le mécanisme moléculaire de cette étape du cycle viral est connu (Craigie, 2001; Delelis et al., 2008), ce qui fait de l'intégrase une cible thérapeutique majeure (Dubey et al., 2007). Il a par ailleurs été montré que cette protéine est également à l'origine de la sélection du site d'intégration du génome viral qui, pour le VIH, correspond essentiellement à des zones activement transcrites (Kvaratskhelia et al., 2014).

Suite à l'intégration du provirus dans le génome cellulaire, celui-ci est transcrit via l'activation du promoteur localisé dans les LTR par des facteurs cellulaires. Cette phase d'expression de l'ADN viral a lieu en deux temps. Tout d'abord, les ARN viraux poly-épissés sont exportés hors du noyau et traduits pour permettre la production des protéines virales Tat, Rev et Nef. Tat va ensuite stimuler la transcription des gènes viraux en interagissant avec la structure secondaire TAR et en recrutant le Complexe de Super Elongation (SEC) au niveau de l'ARN polymérase (Huasong et al., 2013). Parallèlement, Rev va permettre l'export nucléaire des ARN viraux non épissés et mono-épissés. La persistance dans le génome de certaines cellules de provirus non transcrits, phénomène portant le nom de latence virale, est un des obstacles majeurs à la mise en place de traitements permettant de guérir les patients séropositifs (Van Lint et al., 2013).

Une fois exportés hors du noyau, les ARN viraux sont traduits dans le cytoplasme pour donner les différentes protéines virales. La protéine d'enveloppe s'insère dans la membrane plasmique mais c'est le précurseur Gag qui est l'acteur clé du processus de bourgeonnement des particules virales. Cette protéine interagit en effet avec la membrane plasmique, avec des protéines virales, avec l'ARN génome viral et avec certaines protéines cellulaires pour permettre le bourgeonnement de virions depuis la membrane plasmique. Ce bourgeonnement, ayant lieu au niveau de zones membranaires à la composition lipidiques particulières, les radeaux lipidiques, permet la formation de particules virales immatures qui ne deviendront infectieuses que suite au clivage des précurseurs Gag et Gal-Pol par la protéase virale (Sundquist and Kra, 2012).

Au-delà de ce cycle viral classique, des études de plus en plus nombreuses soulignent l'importance de la transmission cellule-cellule du VIH *in vivo*, sans libération de particules virales dans le milieu extracellulaire. Ce mécanisme pourrait expliquer certaines particularités de la pathogenèse du VIH (Costiniuk and Jenabian, 2014).

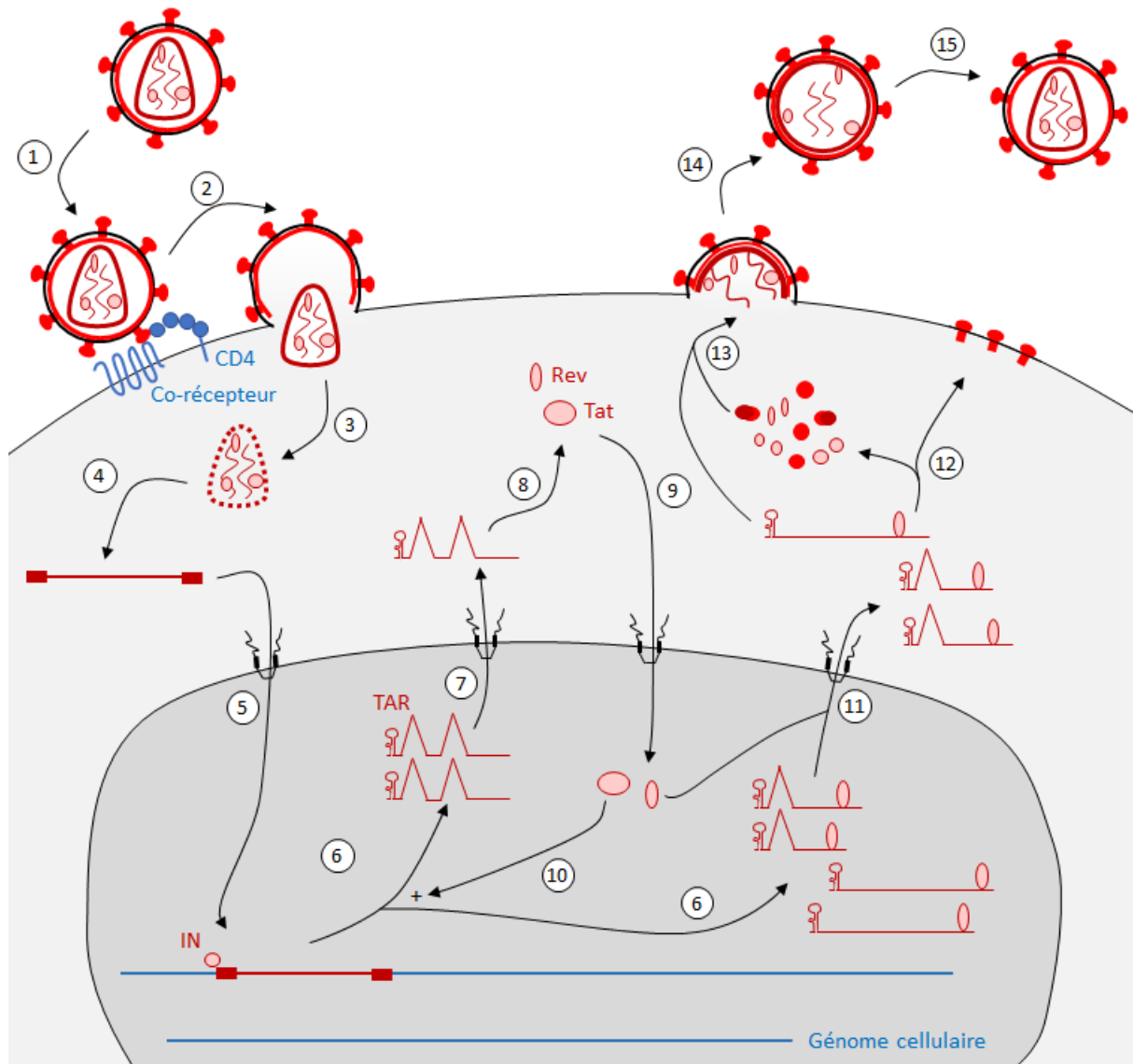


Figure 23 : schéma du cycle de réplication du VIH. 1 : reconnaissance du récepteur CD4 et d'un co-récepteur cellulaire par la protéine d'enveloppe virale. 2 : fusion des membranes virale et cellulaire et libération de la capside virale dans le cytoplasme. 3 : décapsidation. 4 : transcription inverse de l'ARN viral en ADN viral. 5 : import nucléaire de l'ADN viral et insertion dans le génome cellulaire. 6 : transcription du génome viral. 7 : export des ARNm viraux épissés. 8 : traduction des protéines Tat et Rev. 9 : import nucléaire de Tat et Rev. 10 : activation de la transcription virale par Tat. 11 : export des ARN viraux non épissés et mono épissés via Rev. 12 : traduction des autres protéines virales. 13 : assemblage des particules virales au niveau de la membrane plasmique. 14 : bourgeonnement. 15 : maturation des particules par la protéase virale.

b/ Utilisation de facteurs cellulaires facilitateurs

La réplication des virus est un processus complexe faisant appel à différents mécanismes impliquant un grand nombre de protéines. Or, la taille limitée des génomes viraux ne leur permet pas de coder pour l'ensemble des gènes nécessaires à leur propre réplication. Les virus sont ainsi dépendants des cellules qu'ils infectent pour se répliquer. Ils ont évolué pour utiliser au mieux différents facteurs cellulaires facilitant leur multiplication et le VIH ne fait pas exception.

Tout d'abord, le VIH exploite les ressources énergétiques et les voies de synthèse des biomolécules des cellules infectées. En effet, l'infection conduit à la production de particules virales

contenant de l'ARN, des protéines et des lipides : les précurseurs correspondants (NTP, acides aminés et acétyl-CoA) sont ainsi indirectement utilisés par le virus. De plus, l'étape de transcription inverse de l'ARN viral en ADN nécessite l'apport de dNTP cellulaires.

Parallèlement, le cycle de réplication viral comporte de nombreuses étapes nécessitant de l'énergie. Ainsi, en plus de récupérer les produits des voies anaboliques dans la cellule, le VIH utilise l'énergie produite par les voies cataboliques.

Si la dépendance du virus vis-à-vis du métabolisme de la cellule est évidente, peu d'études ont permis de montrer que le VIH pouvait moduler ce métabolisme. Des effets directs de l'infection par le VIH ont été observés sur le métabolisme des lipides sans que les mécanismes précis ne soient élucidés (Rasheed et al., 2008) et il a été montré que la protéine Vpr a un effet direct sur PPAR γ et le récepteur des glucocorticoïdes, perturbant ainsi le fonctionnement du tissu adipeux (Agarwal et al., 2013). Chez les patients séropositifs, des perturbations du métabolisme lipidique mais aussi du métabolisme des glucides ont été observées à l'échelle systémique, elles sont vraisemblablement dues à la fois à l'infection et aux traitements antirétroviraux (Palmer and Crowe, 2012; Stanley and Grinspoon, 2012).

Concernant le métabolisme des glucides, l'observation de modifications concomitantes à l'infection par le VIH peut être expliquée au niveau cellulaire par le fait que l'activation des lymphocytes, principales cellules cibles du VIH, conduit à un changement de leur fonctionnement métabolique. Lors de la réponse immunitaire, les lymphocytes T passent d'un métabolisme oxydatif du glucose au cours duquel la glycolyse est suivie par le cycle de Krebs et la respiration mitochondriale à un métabolisme glycolytique reposant sur la glycolyse et la fermentation lactique. Cette transition métabolique est régulée par les cytokines et est maintenue tout au long de la réponse immunitaire (Michalek and Rathmell, 2011; Palmer et al., 2015). On peut ainsi expliquer que l'inflammation chronique induite par le VIH perturbe le métabolisme des glucides à l'échelle cellulaire, ce métabolisme semblant lui-même jouer un rôle dans la sensibilité au VIH (Loisel-Meyer et al., 2012).

Les virus sont capables d'utiliser l'énergie et les biomolécules présentes dans les cellules, mais ils sont également capables d'interagir ponctuellement avec des partenaires et des structures cellulaires pour effectuer certaines étapes de leur cycle de réplication.

Le VIH est notamment capable de détourner des récepteurs membranaires pour permettre la libération de sa capsid dans le cytoplasme (Klasse, 2012), de recruter des protéines d'export nucléaire pour permettre l'export des ARN viraux (Neville et al., 1997) mais aussi d'interagir avec des domaines membranaires particuliers pour permettre son bourgeonnement (Ono and Freed, 2001).

Enfin, le VIH est capable de détourner des machineries cellulaires complexes pour permettre l'expression de son génome.

En effet, en plus de posséder des séquences reconnues par des facteurs de transcription au niveau de ses LTR, le VIH code pour la protéine Tat qui est capable d'interagir avec la structure TAR de l'ARN viral et d'y recruter le complexe pTEFb sous forme active (He et al., 2010; Sobhian et al., 2010). Ce complexe peut alors phosphoryler le domaine C-terminal de l'ARN polymérase II au niveau du génome viral et activer sa transcription de façon processive.

Le VIH recrute également la machinerie de traduction pour permettre la synthèse des protéines virales. La structure de certains ARNm viraux étant différente de celle des ARNm cellulaires, des mécanismes alternatifs de traduction sont mis en place. Ainsi, la traduction du gène Env débute via un saut de ribosome, la traduction des précurseurs Gag et Gag-Pol peut être initiée de façon coiffe-dépendante ou via un IRES et la traduction du précurseur Gag-Pol n'a lieu que si le ribosome dérape entre les deux gènes pour reculer d'un nucléotide (Guerrero et al., 2015).

De façon générale le VIH est capable de se multiplier dans les cellules infectées car, à chaque étape du cycle viral, il est capable d'utiliser les facteurs cellulaires qui facilitent sa réplication.

c/ Echappement à l'immunité innée

Il serait néanmoins malvenu de considérer la cellule comme un environnement de réplication accueillant. En effet, en plus des réponses immunitaires innées et acquises mises en place à l'échelle de l'organisme, l'infection par le VIH conduit au déclenchement de voies de réponse immunitaire innée à l'échelle cellulaire (Figure 24).

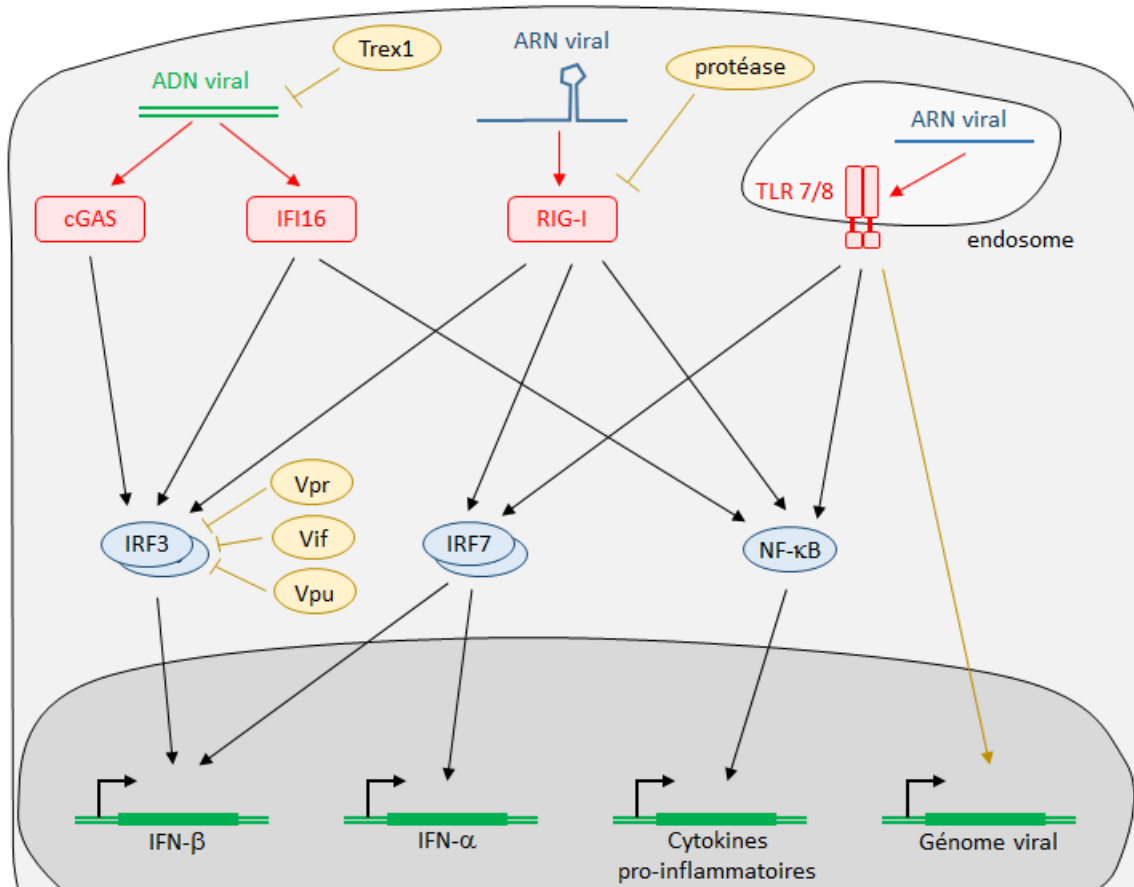


Figure 24 : voies de réponse immunitaire innée cellulaire pouvant être activées par le VIH et mécanismes d'échappement viraux. Le détail des voies n'a pas été représenté. Les senseurs sont indiqués en rouge et les facteurs de transcription en bleu. cGAS et IFI16 reconnaissent l'ADN double brin cytoplasmique, RIG-I reconnaît l'ARN double brin et donc les structures secondaires de l'ARN cytoplasmique et les TLR7 et 8 reconnaissent l'ARN simple brin endosomal. Les façons dont le VIH peut contrecarrer ou détourner ces voies de signalisation sont représentées en jaune.

De façon générale les mécanismes de l'immunité innée cellulaire reposent sur la détection des agents pathogènes par des récepteurs appelés PRR (Pattern-Recognition Receptors) capables d'interagir avec des motifs moléculaires propres aux pathogènes (ou PAMP : Pathogen-Associated Molecular Patterns). Cette reconnaissance permet d'activer différentes voies de signalisation pouvant conduire à l'expression de cytokines voire à la mort cellulaire. Deux types de PRR, induisant une réponse antivirale, sont impliqués dans la reconnaissance du VIH.

Tout d'abord l'hélicase RIG-I est capable de reconnaître des structures secondaires de l'ARN génomique viral. Dans les cellules mononuclées du sang périphérique (PBMC) et dans les macrophages, cette détection induit une faible expression d'interféron (IFN) de type I et de $\text{TNF}\alpha$ permettant l'expression des gènes stimulés par l'IFN (ISG ou Interferon Stimulated Genes) (Berg et al., 2012).

Par ailleurs plusieurs récepteurs de la famille des Toll-like récepteurs (TLR) sont capables de reconnaître l'ARN viral dans certains types cellulaires. Dans les cellules dendritiques, cette détection

par le TLR7 induit la production d'INF α et de TNF α (Beignon et al., 2005; Lepelley et al., 2011). Dans les monocytes, le TLR7 et le TLR8 peuvent reconnaître l'ARN viral. Plusieurs équipes ont montré que cette détection induisait l'activation de l'inflammasome (Chattergoon et al., 2014; Guo et al., 2014) et la sécrétion de cytokines, mais les cytokines identifiées sont différentes selon les études (IL6 et TNF α (Meier et al., 2007), IL β (Guo et al., 2014) ou absence d'IFN (Chattergoon et al., 2014)).

Il a également été montré qu'un censeur n'appartenant pas à la famille des PRR peut reconnaître le VIH. Il s'agit de l'enzyme cGAS, qui est un censeur d'ADN cytoplasmique, et qui est notamment capable de détecter l'ADN viral si celui-ci n'est pas efficacement masqué par la capsid au cours de la transcription inverse. Cette reconnaissance induit une réponse IFN de type 1 (Gao et al., 2013; Lahaye et al., 2013). La protéine cellulaire IFI16, autre senseur de l'ADN cytoplasmique, est également capable d'induire une réponse immunitaire innée après détection de l'ADN viral (Jakobsen et al., 2013).

Seuls les virus capables de dépasser ces blocages peuvent achever leur cycle de réplication. C'est pourquoi le VIH évolue pour contrecarrer ces facteurs de l'immunité innée ou les utiliser au profit de sa multiplication (Figure 24). Il a par exemple été montré que dans certaines cellules dendritiques l'activation du TLR8 améliore la réplication du VIH en augmentant la transcription virale (Gringhuis et al., 2010).

Les protéines virales peuvent également bloquer la signalisation des voies de l'immunité cellulaire en inhibant les intermédiaires impliqués. Ainsi, la protéase virale est capable de séquestrer le détecteur RIG-I (Solis et al., 2011) et les protéines accessoires Vpr et Vif peuvent induire la dégradation de l'intermédiaire IRF3 (Okumura et al., 2008). Cette molécule est également ciblée par la protéine virale Vpu mais le mécanisme mis en jeu est sujet à controverse. Une équipe avait postulé que Vpu induisait la dégradation de IRF3 (Doehle et al., 2012), ce qui a été contredit par la suite (Hotter et al., 2013) avant que d'autres chercheurs montrent que Vpu induit le clivage d'IRF3 par les caspases mais pas sa dégradation complète (Park et al., 2014).

Enfin, certains mécanismes cellulaires limitant le déclenchement de la réponse innée en l'absence de pathogènes sont également mis à profit par le VIH. On peut notamment citer la nucléase Trex1 qui est capable de dégrader les ADN cytoplasmiques. Cette dernière peut cibler les produits de transcription inverse abortive du VIH, ce qui limite la détection du virus par les senseurs cellulaires (Yan et al., 2010).

3/ VIH et facteurs de restriction

a/ Caractéristiques et fonctions des facteurs de restrictions

En plus de devoir échapper aux mécanismes de l'immunité innée, les rétrovirus et notamment le VIH doivent contrecarrer des facteurs de restriction. Ces protéines cellulaires sont définies comme étant capables de bloquer la réplication du virus en agissant sur une étape précise du cycle viral. Elles sont généralement contrecarrées par les virus et les gènes correspondants peuvent être induits par l'IFN et portent des marques de sélection positive. En pratique, cela signifie que le taux de mutations non synonymes observé dans ces gènes ou dans certaines régions de ces gènes est supérieur au taux de mutations synonymes, lui-même considéré comme relevant d'une apparition aléatoire de mutations non sélectionnées par la suite.

A l'heure actuelle, plusieurs facteurs de restriction ont été identifiés comme capables de bloquer le cycle viral du VIH. De façon intéressante, pour chacun de ces facteurs, des données suggèrent un rôle dans la modulation de la réponse immunitaire innée. La compréhension globale de l'activité antivirale de ces protéines nécessite donc des études in vivo plus approfondies que les expériences généralement effectuées avec des lignées cellulaires. Plusieurs travaux réalisés chez la souris ouvrent ainsi des perspectives intéressantes sur les interactions entre immunité innée et facteurs de restriction (Rehwinkel, 2014).

Les découvertes rallongent régulièrement la liste des facteurs de restriction du VIH et des virus apparentés. Seuls les quatre facteurs de restriction les mieux caractérisés seront présentés dans ce manuscrit (Figure 25).

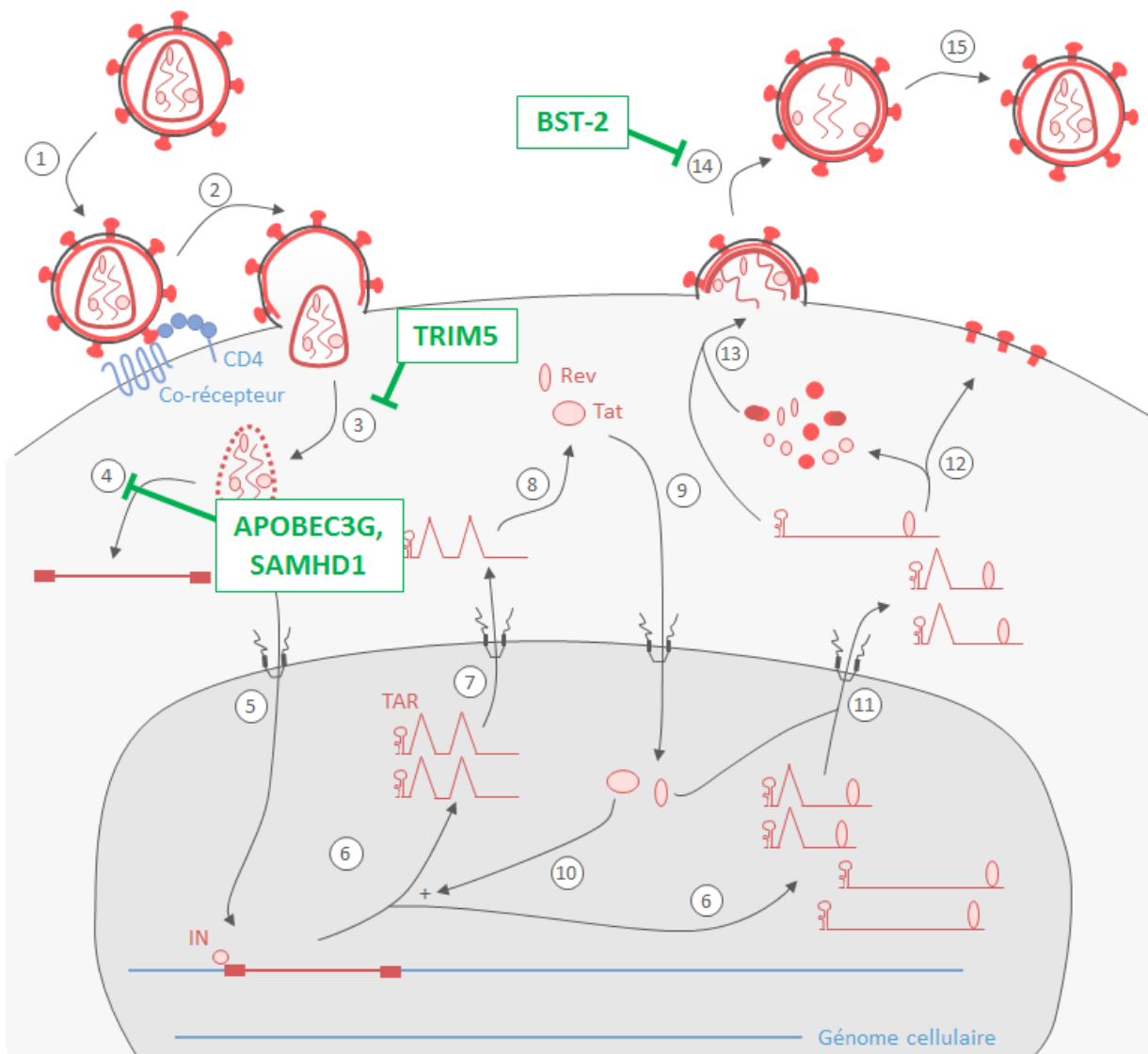


Figure 25 : schéma des étapes du cycle de réplication du VIH bloquées par les quatre facteurs de restriction actuellement identifiés pour ce virus. 3 : décapsidation. 4 : transcription inverse. 14 : bourgeonnement. (numérotation d'après la Figure 23)

Le premier, TRIM5, est particulier car la protéine humaine ne restreint pas le VIH. Néanmoins les protéines TRIM5 d'autres primates sont capables de bloquer ce virus (Sayah et al., 2004; Stremlau et al., 2004). Elles agissent en provoquant une décapsidation précoce qui conduit à l'exposition des acides nucléiques viraux au cours de la transcription inverse (Stremlau et al., 2006). Cependant, le mécanisme sous-jacent reste encore mal compris. En effet, il a été démontré que TRIM5 peut interagir avec la protéine de capsid virale (Stremlau et al., 2006) et que son activité ubiquitine-ligase est nécessaire au blocage du cycle viral avant la transcription inverse (Lienlaf et al., 2011; Pertel et al., 2011). En revanche aucune ubiquitination de la capsid virale par TRIM5 n'a été observée à ce jour. D'autres données expérimentales suggèrent que TRIM5 pourrait également bloquer d'autres étapes du cycle viral (Wu et al., 2006) voire que TRIM5 pourrait être considéré comme un PRR reconnaissant la capsid virale (Pertel et al., 2011).

APOBEC3G est un autre facteur de restriction agissant dans les phases précoces du cycle viral, plus précisément au niveau de la transcription inverse. Ce facteur de restriction a été identifié via l'étude de la protéine virale Vif. En effet, des travaux ont montré que Vif permet l'infection de cellules non permissives par le VIH, vraisemblablement en dégradant un facteur antiviral cellulaire (Madani and Kabat, 1998). C'est en cherchant à identifier ce facteur que des chercheurs ont montré qu'APOBEC3G peut empêcher l'infection par le VIH mais uniquement en l'absence de Vif (Sheehy et al., 2002). Parallèlement, une hypermutation du génome viral a été observée en l'absence de Vif (Lecossier et al., 2003). Il a finalement été démontré que cette hypermutation est due à APOBEC3G (Zhang et al., 2003).

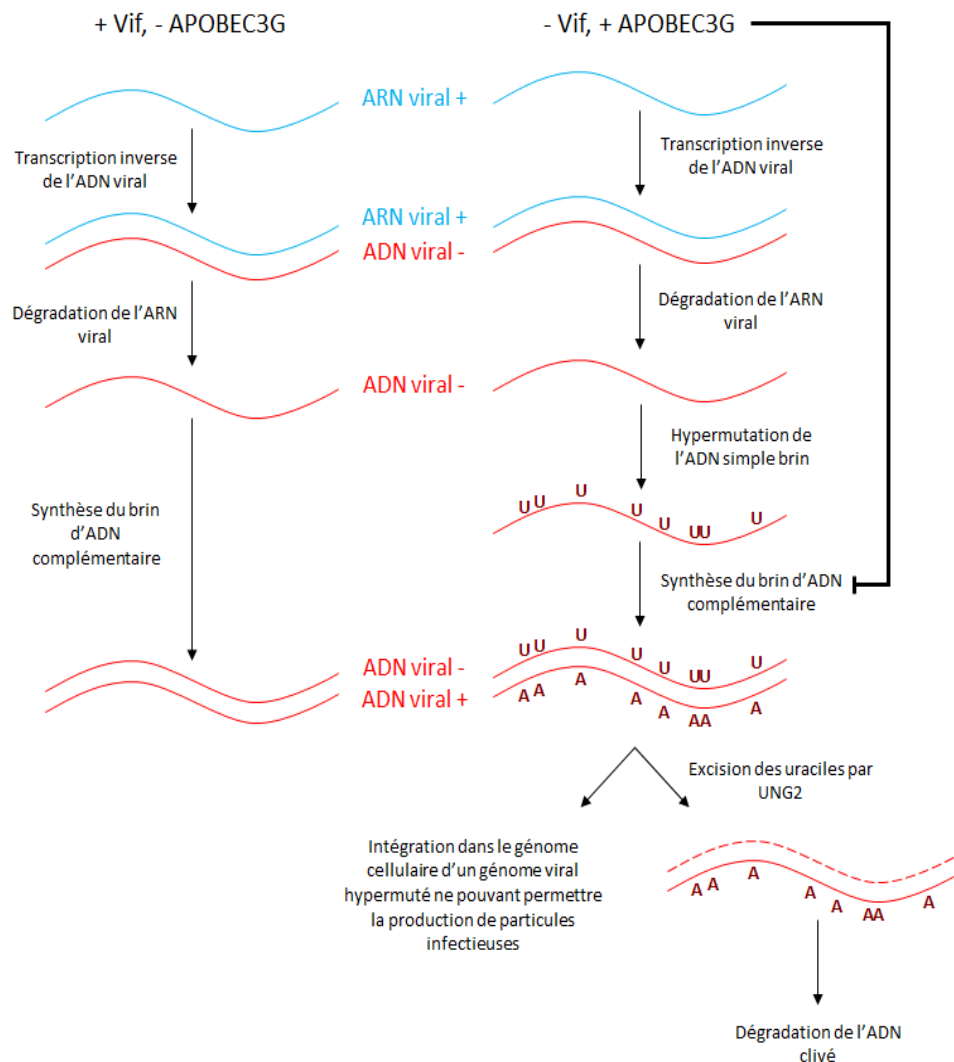


Figure 26 : schéma des effets d'APOBEC3G sur la transcription inverse. En présence d'APOBEC3G, l'ADN viral + subit des mutations C vers U. Ces mutations inhibent la synthèse du brin complémentaire, ce qui est le mécanisme principal de l'arrêt du cycle viral dû à APOBEC3G. Dans les cas où le brin d'ADN complémentaire peut être synthétisé, l'ADN double brin obtenu peut être reconnu par la protéine cellulaire UNG2, qui clive les uraciles dans l'ADN, ce qui conduit à sa dégradation. Il peut également s'intégrer dans le génome cellulaire. Dans ce dernier cas, l'infection virale n'est pas productive car l'hypermutation du génome ne permet pas l'expression de protéines virales fonctionnelles.

En effet, APOBEC3G est une cytidine déaminase capable d'introduire des mutations C (cytidine) vers U (uracile) dans les ADN simple brin. C'est une protéine cytoplasmique qui peut être incorporée dans les particules virales au cours du bourgeonnement. Lorsque ces particules contenant la protéine APOBEC3G infectent une cellule, APOBEC3G introduit des mutations dans l'ADN viral au cours de la transcription inverse. Il a été montré que cette hypermutation du génome viral est délétère

pour le virus (Mangeat et al., 2003) et que l'incorporation d'uraciles dans le premier brin d'ADN viral synthétisé perturbe la transcription du second brin d'ADN (Klarmann et al., 2003). Plusieurs mécanismes sont vraisemblablement impliqués, notamment l'inhibition de la transcription inverse via l'hypermutation du premier brin d'ADN synthétisé (Bishop et al., 2008) (Figure 26).

APOBEC3G n'est pas un facteur de restriction spécifique du VIH, il est capable de bloquer la transcription inverse d'autres rétrovirus (Mangeat et al., 2003). La première démonstration formelle du mécanisme d'action d'APOBEC3G a d'ailleurs été réalisé avec le virus de la leucémie murine (MLV) (Harris et al., 2003).

De plus APOBEC3G appartient à la famille des cytidines déaminases APOBEC3, qui est composée de sept protéines différentes. D'autres membres de cette famille sont capables de bloquer l'infection par le VIH. On peut notamment citer APOBEC3A (Berger et al., 2011b) et APOBEC3F (Zheng et al., 2004b). Il a été montré qu'APOBEC3B est capable de restreindre le VIH et n'est pas contrecarré par Vif, mais cette protéine de la famille des APOBEC n'est pas exprimée dans les cellules cibles du VIH (Doehle et al., 2005).

SAMHD1 est le troisième facteur de restriction connu capable de bloquer le VIH avant l'insertion du génome viral dans le génome cellulaire. Les connaissances actuelles sur cette protéine seront récapitulées plus tard dans ce manuscrit (II/B).

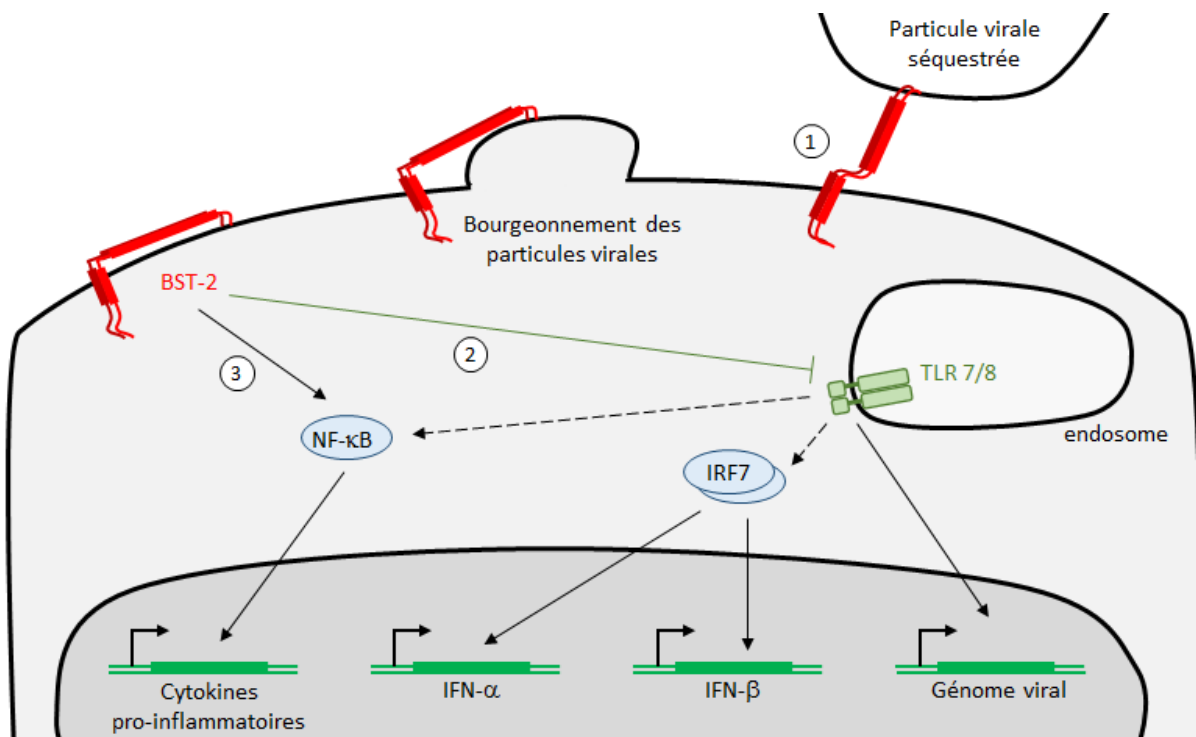


Figure 27 : schéma des mécanismes par lesquels BST-2 inhibe l'infection par le VIH. 1 : séquestration des particules virales au moment de leur bourgeonnement. 2 : inhibition de l'activité des TLR7/8 (effet indépendant de l'infection par le VIH). 3 : activation du facteur de transcription NF-κB.

Enfin, contrairement aux trois autres facteurs de restriction identifiés à ce jour, la tetherine, également appelée BST-2, bloque le cycle viral dans sa phase tardive (Figure 27). Ce facteur de restriction a été identifié comme cible de la protéine virale Vpu et comme limitant le bourgeonnement des particules virales (Neil et al., 2008; VanDamme et al., 2008). Cette capacité de séquestration est due à la structure particulière de BST-2, qui possède à la fois un domaine transmembranaire et une ancre GPI. Cette protéine fonctionne sous forme de dimères, qui possèdent donc chacun quatre sites d'ancrage à la membrane plasmique. Au cours du bourgeonnement viral, la répartition de ces sites d'ancrage entre les membranes cellulaire et virale permet à BST-2 de physiquement séquestrer les virions et d'empêcher leur libération dans le milieu extracellulaire (Perez-caballero et al., 2009). Il est

intéressant de noter que BST-2 est également capable de limiter la transmission cellule-cellule du VIH (Casartelli et al., 2010).

Des chercheurs ont montré que l'activité antivirale de BST-2 n'était pas limitée au VIH, ni même aux rétrovirus. Ce facteur pourrait en fait avoir une activité globale contre les virus possédant une membrane lipidique issue de la membrane plasmique (Jouvenet et al., 2009).

Des données expérimentales montrent que l'activité anti-VIH de BST-2 n'est pas uniquement due à sa capacité de séquestrer les particules virales. Ainsi, plusieurs équipes ont montré que BST-2 pouvait agir comme un PRR et activer le facteur de transcription NF- κ B lors de l'infection (Cocka and Bates, 2012; Galão et al., 2012; Tokarev et al., 2013). BST-2 est également capable de moduler la signalisation des TLR7 et 9 dans certaines cellules dendritiques (Cao et al., 2009). Enfin, il a récemment été montré que BST-2 peut inhiber l'expression transitoire des protéines (Narkpuk et al., 2014) et l'effet de cette propriété sur l'infection par le VIH n'a pas encore été élucidé.

b/ Mécanismes d'échappement viral

Le fait d'être contrecarrés par des protéines virales est un des éléments de définition des facteurs de restriction. La plupart d'entre eux ont d'ailleurs été identifiés en cherchant quel facteur cellulaire était bloqué par une protéine virale nécessaire à l'infection de cellules non permissives (Tableau 5).

Le seul facteur de restriction qui n'est pas contrecarré par une protéine accessoire du VIH est TRIM5. En effet, cette protéine n'est pas capable de restreindre l'infection par le VIH chez l'homme. Néanmoins il a été montré que lorsqu'un variant capable de bloquer le VIH est surexprimé dans des cellules infectées par le VIH, le virus s'adapte via des mutations dans le gène codant pour la capsid virale (Pacheco et al., 2010).

APOBEC3G a été identifié comme le facteur de restriction contrecarré par la protéine virale Vif. Assez rapidement après cette identification, il a été montré que Vif agissait à la fois en inhibant la traduction d'APOBEC3G et en promouvant la dégradation de cette protéine de façon dépendante du protéasome (Stopak et al., 2003). La combinaison de ces deux mécanismes permet à Vif d'empêcher l'incorporation d'APOBEC3G dans les particules virales (Stopak et al., 2003). Cette capacité de la protéine virale Vif à bloquer la protéine cellulaire APOBEC3G est dépendante de la souche virale et de l'espèce infectée (Mariani et al., 2003).

Il a été montré que Vif induit la dégradation d'APOBEC3G par le protéasome en interagissant avec cette protéine et avec un complexe E3-ligase de type SCF comprenant la culline 5, les élongines B et C et l'adaptateur Rbx1 (Yu et al., 2003). Vif est également capable d'interagir avec CBF β (Zhou et al., 2012) et cette protéine cellulaire est nécessaire à la dégradation d'APOBEC3G par Vif (Jäger et al., 2011; Zhang et al., 2011).

Vif est capable d'induire la dégradation d'autres protéines de la famille APOBEC3, notamment APOBEC3F, en recrutant le même complexe ubiquitine-ligase (Liu et al., 2005).

De façon intéressante, les différents virus de l'immunodéficience utilisent des protéines différentes pour contrecarrer BST-2 : Vpu pour le VIH-1, Env pour le VIH-2 et Nef pour la plupart des VIS. Cela suggère que ce facteur de restriction peut jouer un rôle important dans la barrière d'espèce et que les VIH-1 et 2 se sont adaptés de façon différente à leur nouvel hôte humain (Sauter et al., 2009).

Concernant le VIH-1, il était connu que la protéine Vpu est capable d'induire la dégradation du récepteur membranaire CD4 en interagissant avec l'ubiquitine-ligase β TrCP (Margottin et al., 1998). Il a par la suite été montré que le domaine d'interaction avec β TrCP de Vpu mais aussi son domaine transmembranaire sont nécessaires à l'inhibition de BST-2 (Schubert et al., 1996). Des données

suggèrent que le domaine transmembranaire de Vpu permet son interaction avec BST-2 (Iwabu et al., 2009).

Néanmoins, bien que le rôle de β TrCP dans le blocage de BST-2 par Vpu ait été démontré simultanément par plusieurs équipes, ces dernières proposent des mécanismes d'action différents. Pour certains chercheurs, Vpu inhibe l'activité antivirale de BST-2 sans modifier la localisation cellulaire de cette protéine ni la dégrader (Miyagi et al., 2009). Pour d'autres, l'activité de Vpu repose majoritairement sur une internalisation de BST-2 (Mitchell et al., 2009). La majorité des publications suggère cependant que Vpu est capable d'induire la dégradation de BST-2. Mais certaines observent une dégradation dépendante des lysosomes (Iwabu et al., 2009; Mitchell et al., 2009) alors que les autres observent une dégradation dépendante du protéasome (Douglas et al., 2009; Goffinet et al., 2009; Mangeat et al., 2009). Le mécanisme exact reste encore indéterminé.

Protéine virale	Mécanisme	Partenaires impliqués	Facteur de restriction
CA	mutation et perte d'interaction		TRIM5
Vif	inhibition de la traduction		APOBEC3G
Vif	induction de l'ubiquitination et de la dégradation	Cul5/EloB/EloC/Rbx1	APOBEC3G
Vpx	induction de l'ubiquitination et de la dégradation	Cul4/DDB1/Dcaf1	SAMHD1
Vpu (VIH-1)	induction de la dégradation/relocalisation ?	β -TrCP	BST-2
Env (VIH-2)	Relocalisation		BST-2

Tableau 5 : récapitulatif des protéines virales capables de contrecarrer les principaux facteurs de restriction des virus de l'immunodéficience humaine.

B/ SAMHD1, du phénotype à la caractérisation moléculaire

1/ Identification d'un nouveau facteur de restriction

a/ Importance de Vpx pour l'infection de certains types cellulaires

En utilisant des virus n'exprimant pas la protéine virale Vpx, il a été montré dès 1996 que cette protéine, exprimée par le VIH-2 et les VIS apparentés, est nécessaire à l'infection des macrophages de macaques par le virus VISsm mais pas à celle des lymphocytes (Fletcher et al., 1996). L'influence de Vpx sur la pathogenèse du VISsm chez le macaque a été étudiée et il s'avère qu'en l'absence de Vpx ce virus est moins pathogène (virémie plasmatique réduite, survie des singes infectés, absence d'amplification locale après infection des muqueuses) (Hirsch et al., 1998). Le VISsm n'exprimant plus la protéine Vpx est ainsi rapidement dominé par la souche sauvage en cas de co-infection : chez des singes inoculés avec des quantités équivalentes des deux virus, moins de deux pourcents des virus circulants quatre jours après infections n'expriment pas Vpx (Hirsch et al., 1998).

D'autres travaux ont permis de constater l'importance de Vpx pour l'infection des cellules myéloïdes chez le macaque et chez l'homme. Ce phénotype a été observé dans les cellules dendritiques dérivées de monocytes (Goujon et al., 2006), dans les monocytes (Goujon et al., 2008; Kaushik et al., 2009) et dans les macrophages (Fletcher et al., 1996; Goujon et al., 2008; Sharova et al., 2008; Srivastava et al., 2008). En revanche Vpx n'est pas nécessaire pour l'infection des lymphocytes (Fletcher et al., 1996; Goujon et al., 2008; Hirsch et al., 1998).

Le rôle de Vpx dans l'infection des cellules myéloïdes a été mis en évidence en utilisant des VIS ou des souches de VIH-2 n'exprimant plus cette protéine, moins infectieux que les virus sauvages (Fletcher et al., 1996; Sharova et al., 2008). Néanmoins il a rapidement été observé que l'apport de la protéine Vpx en trans avant ou pendant l'infection par un virus n'exprimant pas Vpx, comme le VIH-1,

permettait d'augmenter significativement l'infection des cellules myéloïdes par ce virus (Berger et al., 2009; Goujon et al., 2006; Kaushik et al., 2009).

La protéine Vpx du VIH-2 et de certains VIS est capable de faciliter l'infection des cellules myéloïdes par le VIH-1 mais également par d'autres rétrovirus comme le Virus de l'Immunodéficience Féline (Goujon et al., 2007). Par ailleurs, les protéines Vpx des différents VIS de la lignée du VIH-2 ne sont pas équivalentes vis-à-vis de l'infection des cellules myéloïdes. En effet, les Vpx du VISmac et du VIH-2 permettent de restaurer l'infection des cellules dendritiques par le VISmac delta Vpx, mais pas le Vpx du VISrcm (Tableau6) (Goujon et al., 2007).

Des études ont montré que Vpx permet d'augmenter l'infection des cellules myéloïdes en agissant sur les phases précoces du cycle viral. Ceci est cohérent avec le fait que Vpx soit internalisé dans les particules virales (Selig et al., 1999) mais l'étape précisément ciblée est sujette à discussion. En effet, certaines équipes ont mis en évidence un effet dès les premières étapes de la transcription inverse (Srivastava et al., 2008), d'autres n'observent un effet que sur la formation des cercles à deux LTR (marqueurs d'import nucléaire) (Fletcher et al., 1996) et d'autres observent un effet sur les étapes précoces ou les étapes tardives de la transcription inverse en fonction des virus considérés (Goujon et al., 2007). Il a été montré que le Vpx d'une souche de VISsm est capable d'interagir avec la machinerie de transcription inverse du virus (Fletcher et al., 1996).

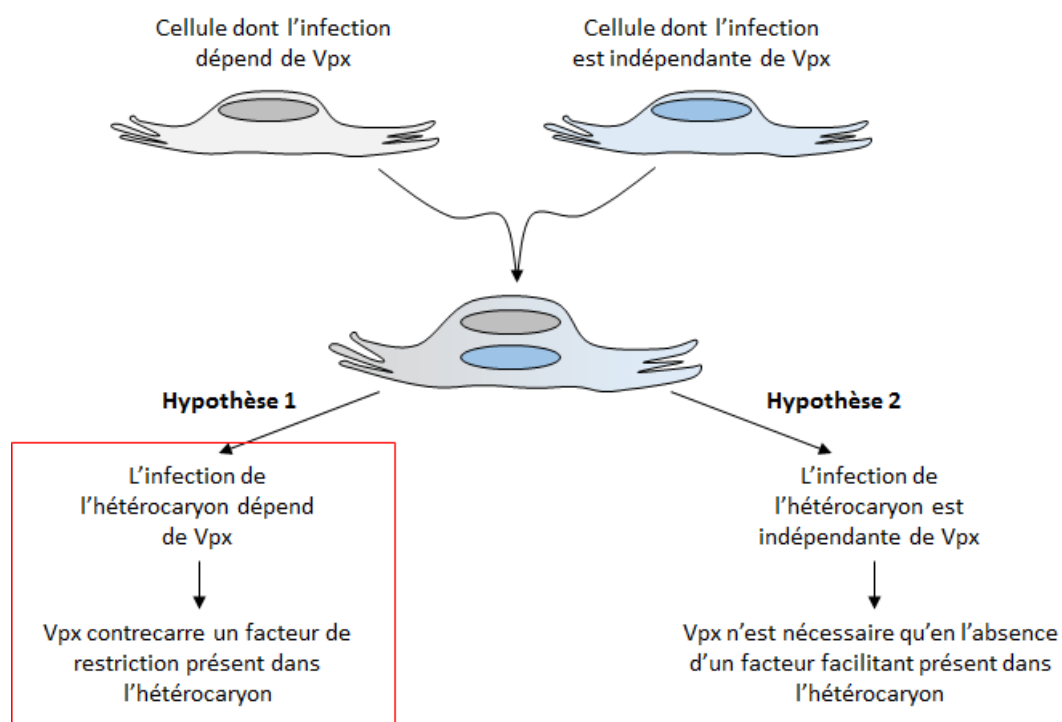


Figure 28 : principe des expériences de mesure d'infectivité d'hétérocaryons.

Des expériences de formation d'hétérocaryons issus de la fusion de cellules dont l'infection est modulée par Vpx avec des cellules dont l'infection est indépendante de Vpx ont permis de montrer que Vpx agit en contrecarrant un facteur de restriction cellulaire (Figure 28). En effet, les hétérocaryons issus de la fusion de macrophages avec des cellules COS permissives ne sont pas sensibles à l'infection par le VISsm (Sharova et al., 2008) et des hétérocaryons issus de la fusion de monocytes et de cellules HeLa permissives ne sont sensibles au VIH-1 que si les cellules HeLa surexpriment la protéine Vpx (Kaushik et al., 2009). Cela montre que Vpx est nécessaire pour contrecarrer un facteur initialement absent des cellules permissives mais apporté par les cellules résistantes lors de la formation des hétérocaryons.

Avant même l'identification de ce facteur, des résultats préliminaires sur le mécanisme mis en œuvre par Vpx rappelaient le mécanisme utilisé par Vif pour contrecarrer APOBEC3G. En effet, il a été montré que l'activité du protéasome est nécessaire pour que Vpx permette d'améliorer l'infection des macrophages (Sharova et al., 2008). De plus, Vpx du VISmac239 interagit avec DDB1 et Dcaf1, qui sont des sous-unités du complexe E3-ligase basé sur la culline 4A (CRL4), ainsi qu'avec DDA1 qui est un régulateur de ce complexe (Srivastava et al., 2008). Il a également été montré que les Vpx des VISrcm et sm ainsi que du VIH-2 sont capables d'interagir avec différentes sous-unités du CRL4 (Goujon et al., 2008; Sharova et al., 2008; Srivastava et al., 2008).

En revanche, l'importance de l'interaction entre Vpx et le CRL4 pour la facilitation de l'infection des cellules myéloïdes par Vpx dépend des études et des types cellulaires. Dans les cellules dendritiques, la capacité d'un Vpx donné apporté en trans à augmenter l'infection par le VIH-1 n'est pas liée à sa capacité à interagir avec Dcaf1. Ainsi, le Vpx du VISrcm interagit avec Dcaf1 mais n'augmente pas l'infection des cellules dendritiques par le VIH-1 alors que le Vpx du VIH-2 augmente l'infection par le VIH-1 mais interagit très faiblement avec Dcaf1 (Tableau 6) (Goujon et al., 2007). De façon comparable, dans les cellules THP-1 différenciées, l'extinction de l'expression de Dcaf1 ou de DDB1 par des shARN n'empêche pas la protéine Vpx du VISmac d'augmenter l'infection par le virus correspondant ou par le VIH-1 (Goujon et al., 2007).

Souche de Vpx	Interaction avec Dcaf1	Augmentation de l'infection des cellules dendritiques par le VIH-1
Mac	Forte	Oui
VIH-2	Très faible	Oui
Rcm	Forte	Non

Tableau 6 : récapitulatif des observations concernant la capacité de différents Vpx à interagir avec Dcaf1 et à augmenter l'infection des cellules dendritiques par le VIH-1.

A l'inverse, dans les macrophages, l'extinction de l'expression de DDB1 diminue fortement l'infection par le VISsm et empêche la protéine Vpx du VISsm d'augmenter l'infection par le VIH-1 (Sharova et al., 2008). L'utilisation de mutants de Vpx du VISmac et du VIH-2 incapables d'interagir avec Dcaf1 a permis de confirmer l'importance de cette interaction dans l'infection des macrophages par les virus correspondants et de montrer que la perte de cette interaction affecte la transcription inverse (Srivastava et al., 2008).

L'ensemble de ces données montre que la protéine virale Vpx est capable de contrecarrer un facteur de restriction cellulaire actif dans les cellules myéloïdes mais pas dans les lymphocytes. Ce facteur de restriction agit pendant les phases précoces du cycle viral, entre le début de la transcription inverse et l'import nucléaire de l'ADN viral. Vpx contrecarre vraisemblablement ce facteur cellulaire en recrutant une ubiquitine ligase pour induire sa dégradation par le protéasome. Dans certains types cellulaires, notamment les macrophages, cette ubiquitine ligase est le complexe CRL4.

b/ Caractérisation de l'activité anti-VIH de SAMHD1

Le facteur de restriction contrecarré par Vpx a été identifié indépendamment par trois équipes en 2011 : il s'agit de la protéine SAMHD1 (Berger et al., 2011a; Hrecka et al., 2011; Laguette et al., 2011). Dans les trois études, SAMHD1 a été identifié à partir de techniques permettant de purifier les protéines interagissant avec Vpx, qu'il s'agisse de purifier les partenaires du Vpx du VISmac251 dans des cellules THP-1 (Laguette et al., 2011), ceux du Vpx du VISmac239 et du complexe CRL4 dans des cellules 293T (Hrecka et al., 2011) ou ceux du Vpx du VISsm dans des cellules 293T (Berger et al., 2011a).

Il a été montré que SAMHD1 est exprimé dans les cellules de la lignée myéloïde et dans la lignée cellulaire THP-1 mais pas dans différentes lignées permissives à l'infection par le VIH-1 (Laguette et al., 2011). SAMHD1 interagit avec la protéine Vpx et cette dernière induit sa dégradation par le protéasome (Berger et al., 2011a; Hrecka et al., 2011; Laguette et al., 2011). De façon intéressante cette dernière propriété n'est pas partagée par l'ensemble des Vpx puisque la protéine Vpx des VISrcm, déjà identifiée comme incapable d'induire l'infection des cellules dendritiques par le VIH-1 (Goujon et al., 2007), n'induit pas la dégradation de SAMHD1 (Laguette et al., 2011). Il a également été confirmé que, dans les macrophages, la dégradation de SAMHD1 est induite par Vpx via le recrutement du complexe CRL4 (Hrecka et al., 2011).

L'effet de SAMHD1 sur l'infection par le VIH-1 a été étudié en surexprimant cette protéine ou en bloquant son expression dans différents types cellulaires. Il a ainsi été montré que SAMHD1 restreint l'infection par le VIH-1 dans les cellules THP-1, dans les cellules dendritiques (Laguette et al., 2011), dans les macrophages (Hrecka et al., 2011) et dans les monocytes (Berger et al., 2011a).

Par ailleurs il a été observé que l'expression de SAMHD1 est induite par l'IFN α dans les monocytes (Berger et al., 2011a), ce qui est fréquemment le cas des facteurs antiviraux.

c/ Une activité plus étendue qu'on ne le pensait

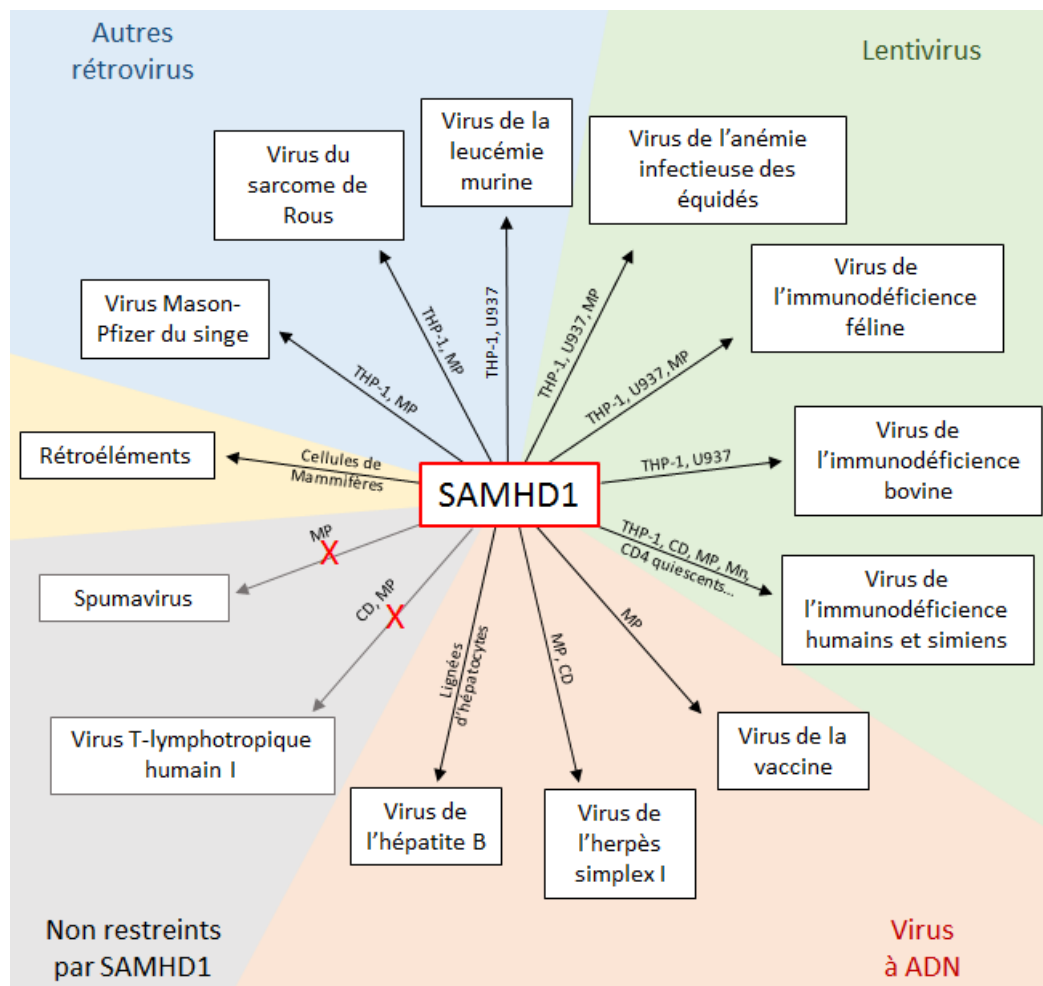


Figure 29 : schéma récapitulant les virus contre lesquels l'activité antivirale de SAMHD1 a été testée. Les types cellulaires correspondants sont indiqués. La liste n'en est pas exhaustive pour les virus de l'immunodéficience humains et simiens. CD : cellules dendritiques. MP : macrophages. Mn : monocytes.

SAMHD1 a été initialement identifié comme un facteur de restriction du VIH-1 dans les cellules myéloïdes pouvant être contrecarré par la protéine virale Vpx. En réalité, ce facteur est actif dans d'autres types cellulaires et contre d'autres virus (Figure 29).

En utilisant le même type d'expériences que celles mises en œuvre dans les cellules myéloïdes, deux équipes ont montré que SAMHD1 limite l'infection des lymphocytes CD4⁺ quiescents par le VIH-1 (Baldauf et al., 2012; Descours et al., 2012). En effet, SAMHD1 est exprimé dans ces cellules et un traitement par la protéine Vpx induit sa dégradation et augmente l'infection par le VIH-1. Par ailleurs, des lymphocytes CD4⁺ quiescents n'exprimant pas SAMHD1 sont plus sensibles à l'infection par le VIH-1, que cette expression réduite de SAMHD1 soit due à une anomalie génétique des donneurs ou à un traitement par ARN interférence.

Il a par ailleurs été montré que SAMHD1 restreint l'infection par le VIH-1 dans une partie des lymphocytes CD4⁺ mémoires (Tabler et al., 2014). Les lymphocytes mémoires sont généralement classés en quatre grandes catégories en fonction des marqueurs qu'ils expriment : les naïfs (T_N), les lymphocytes à mémoire centrale (T_{CM}), les effecteurs (T_{EM}) et les effecteurs terminaux (T_{TE}). Parmi les T_{CM}, une population possédant des propriétés de cellules souches a été identifiée : les T_{SCM}. Une partie de ces cellules exprime les récepteurs du VIH-1, mais leur infection par ce virus est abortive dans plus de 90% des cas. Or ces cellules expriment SAMHD1 et la dégradation de cette protéine via Vpx augmente fortement la fréquence d'infection productive par le VIH-1 (Tabler et al., 2014).

Ainsi, contrairement à la théorie exprimée au moment de son identification, SAMHD1 n'est pas un facteur de restriction spécifique des cellules myéloïdes. Il bloque également l'infection de certains types de lymphocytes par le VIH-1.

Par ailleurs SAMHD1 n'est pas un facteur de restriction spécifique des virus de l'immunodéficience humains et simiens.

Il a été montré que cette protéine est également capable de limiter l'infection de cellules U937 ou THP-1 par les virus de l'immunodéficience bovin (VIB) et félin (VIF) ainsi que d'autres rétrovirus comme les virus de la leucémie murine B ou N tropiques (MLV) et le virus de l'anémie infectieuse des équidés (EIAV) (White et al., 2013a). Des résultats similaires pour le VIF et le EIAV ont été obtenus dans des macrophages (Gramberg et al., 2013). En revanche, dans le cas du MLV, si le traitement des macrophages par la protéine Vpx permet d'augmenter l'efficacité de la transcription inverse, un second blocage plus tardif empêche l'infection productive de ces cellules (Gramberg et al., 2013). Ces résultats obtenus pour le VIF et le MLV sont en accord avec des observations antérieures à l'identification de SAMHD1 comme cible de Vpx (Goujon et al., 2007). Il a également été démontré que SAMHD1 empêche l'infection des macrophages et des THP-1 par le virus du sarcome de Rous et par le MPMV (Gramberg et al., 2013). Cela fait de SAMHD1 un facteur de restriction actif contre un certain nombre de rétrovirus, appartenant à différents genres.

Néanmoins, SAMHD1 n'est pas actif contre l'ensemble des rétrovirus. En effet, le traitement de macrophages par Vpx ne permet pas d'augmenter leur infection par la souche prototype des spumavirus, ce qui pourrait être lié au fait que la transcription inverse des spumavirus a lieu avant leur entrée dans la cellule cible (Gramberg et al., 2013). En ce qui concerne le virus HTLV-1, le traitement des cellules dendritiques et des macrophages par la protéine virale Vpx avant leur co-culture avec des cellules Jurkat infectieuses ne permet pas d'augmenter l'infection de ces cellules primaires (Gramberg et al., 2013). Il a cependant été montré que SAMHD1 a un effet lors de l'infection des monocytes par HTLV-1 en induisant l'apoptose de ces cellules (Sze et al., 2013).

Enfin SAMHD1 a un effet inhibiteur sur la rétro-transposition des éléments LINEs et cette propriété de SAMHD1 est conservée chez l'ensemble des Mammifères pour laquelle elle a été testée (Zhao et al., 2013). SAMHD1 est donc un facteur de restriction actif contre un certain nombre de rétrovirus mais également contre des rétro-éléments endogènes des Mammifères. Ceci suggère que cette protéine est importante pour le fonctionnement de la cellule même en dehors des événements d'infection virale.

Il s'avère d'autre part que SAMHD1 peut contrecarrer d'autres virus que les rétrovirus, notamment des virus à ADN.

Il a notamment été montré que SAMHD1 limite l'infection des macrophages par le virus de la vaccine et celle des THP-1, des macrophages et des cellules dendritiques par le HSV-1 (Hollenbaugh et al., 2013). Concernant le HSV-1, cet effet a été confirmé et intervient après l'expression des gènes viraux précoces (Kim et al., 2013). SAMHD1 a également été caractérisé comme capable de limiter l'infection de lignées cellulaires hépatiques par le virus de l'hépatite B et, comme précédemment décrit pour les monocytes (Berger et al., 2011a), le traitement de ces lignées par de l'IFN α induit l'expression de SAMHD1 (Chen et al., 2014).

SAMHD1 est ainsi capable de bloquer différents virus à ADN susceptibles d'infecter l'Homme. L'étude du variant porcine de cette protéine a permis de montrer qu'il est capable de limiter l'infection de la lignée cellulaire porcine MARC-145 par le virus du syndrome reproducteur et respiratoire porcine (PRRSV) (Yang et al., 2014b). SAMHD1 est donc capable d'agir contre un virus à ARN simple brin de polarité positive ne faisant pas partie des rétrovirus.

2/ Caractérisation de deux activités catalytiques

SAMHD1 est une protéine de 72kDa qui possède deux domaines caractéristiques. D'une part un domaine SAM (Sterile Alpha Motif), qui est susceptible d'être impliqué dans des interactions protéine-protéine, qu'il s'agisse d'homo-multimérisation, d'hétéro-multimérisation ou d'interaction avec des protéines sans domaine SAM, mais pouvant également être impliqué dans des interactions avec des acides nucléiques, notamment des ARN (Kim and Bowie, 2003). D'autre part un domaine HD (Hydrolase, Histidine/Aspartate), qui contient une histidine et un acide aspartique et est susceptible de posséder une activité phosphohydrolase (Aravind and Koonin, 1998). Lors de l'identification de SAMHD1 en tant que facteur de restriction du VIH, aucune activité catalytique de cette protéine n'avait été caractérisée. Deux activités distinctes ont depuis été identifiées.

α / SAMHD1 est une dNTP triphosphohydrolase

Deux études indépendantes ont permis de montrer que SAMHD1 possède une activité désoxynucléosides tri-phosphates (dNTP) triphosphohydrolase, c'est-à-dire est capable de cliver le groupe triphosphate des dNTP (Goldstone et al., 2011; Powell et al., 2011). Ce clivage direct sans libération progressive de mono-phosphate n'avait jamais été caractérisé chez les eucaryotes auparavant.

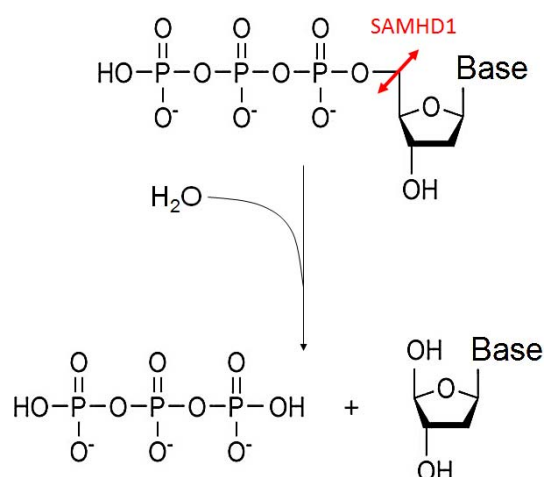


Figure 30 : site de clivage des dNTP par SAMHD1

Cette activité est dépendante d'ions métalliques, essentiellement Mg^{2+} et Mn^{2+} et dépend uniquement du domaine HD de la protéine (Powell et al., 2011). Elle cible précisément la liaison phosphoester entre le phosphate alpha et le désoxyribose des dNTP (Figure 30). La vitesse de dégradation du dGTP par SAMHD1 est environ vingt fois plus élevée que celle de la dégradation des autres dNTP et ces derniers ne sont clivés efficacement par SAMHD1 qu'en présence de dGTP. Par ailleurs, si la vitesse maximale de clivage par SAMHD1 est identique pour le dCTP, le dATP et le dTTP, le dCTP est clivé moins efficacement que les autres dNTP à concentration de dGTP et donc niveau d'activation de SAMHD1 équivalent (Powell et al., 2011).

L'activité de SAMHD1 a été testée sur d'autres types de nucléotides et cette protéine n'est pas active contre les désoxynucléosides mono ou di-phosphates (Powell et al., 2011). De plus, SAMHD1 ne semble pas cliver les nucléosides tri-phosphates, même si une des deux études a mis en évidence une faible activité sur le GTP (Powell et al., 2011). Enfin, aucune activité n'a été observée sur des acides nucléiques polymérisés, qu'il s'agisse d'ADN double brin, d'ADN simple brin ou d'ARN transcrit in vitro (Goldstone et al., 2011).

SAMHD1 est donc une dNTP triphosphohydrolase activée par le dGTP. Cette activité a été observée in vivo après avoir été caractérisée in vitro (Kim et al., 2012). Plusieurs équipes ont élucidé la structure de monomères et de multimères de SAMHD1, ce qui a permis de mieux caractériser la façon dont son activité dNTPase est régulée.

Les premières études avaient montré que SAMHD1 est activé par le dGTP mais pas par ses équivalents di, mono ou sans phosphate (Powell et al., 2011) et peut également être activé par le GTP (Hansen et al., 2014) et dans une moindre mesure par ses équivalents di ou mono-phosphate (Amie et al., 2013). A l'heure actuelle il a été montré que, si SAMHD1 est majoritairement sous forme monomérique ou dimérique dans les cellules, seule sa forme tétramérique possède une activité dNTPase (Ji et al., 2013; Yan et al., 2013; Zhu et al., 2013). La formation de ces tétramères est régulée de façon allostérique et séquentielle. Tout d'abord SAMHD1 interagit avec le GTP, ce qui promeut la formation de dimères. L'interaction avec le dGTP peut conduire au même résultat mais de façon moins efficace. La reconnaissance par ces dimères d'un dNTP, quel qu'il soit, au niveau d'un site régulateur promeut la formation de tétramères possédant une activité dNTPase (Hansen et al., 2014; Ji et al., 2014; Koharudin et al., 2014; Miazzi et al., 2014) (Figure 31). Cette forme activée de SAMHD1 est capable de cliver les dNTP se fixant au niveau des sites catalytiques et persiste après la diminution de la concentration cellulaire des dNTP et du GTP activateurs (Hansen et al., 2014).

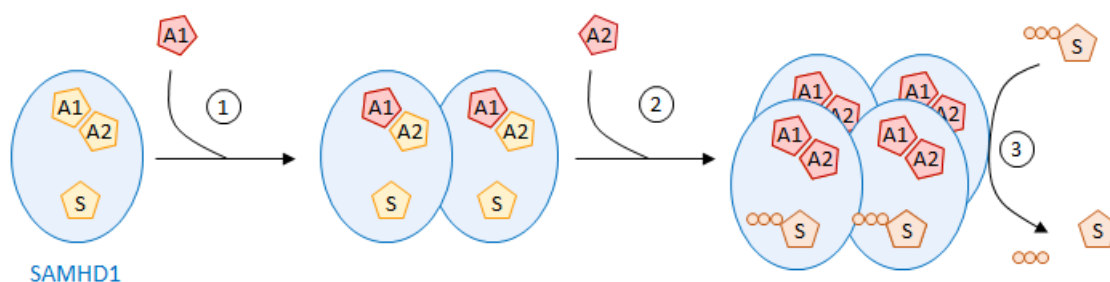


Figure 31 : schéma de l'activation séquentielle l'activité dNTPase de SAMHD1. 1 : fixation d'un GTP ou d'un dGTP au niveau du site activateur 1 (A1) induisant la dimérisation de la protéine. 2 : fixation d'un dNTP au niveau des sites activateurs 2 (A2) du dimère induisant la tétramérisation de la protéine. 3 : activité catalytique de ce tétramère sur les dNTP substrats reconnaissant le site catalytique (S) des enzymes et clivage du groupe triphosphate.

b/ SAMHD1 est une exonucléase

Il a par la suite été montré que SAMHD1 peut interagir avec des acides nucléiques in vitro, avec une préférence pour l'ARN par rapport à l'ADN. Cette interaction dépend d'un domaine comprenant le domaine HD mais est indépendante du domaine SAM. En revanche aucune activité nucléase de SAMHD1 n'a été mise en évidence dans cette étude (Goncalves et al., 2012).

Il a été confirmé que SAMHD1 interagit avec l'ARN simple brin et l'ADN simple brin *in vivo*, en ayant une préférence pour l'ARN, mais pas avec les acides nucléiques doubles brins, y compris les hétéroduplexes. *In vivo* non plus cette interaction ne dépend pas du domaine SAM (Tüngler et al., 2013).

Une autre équipe, ayant réalisé des expériences *in vitro* en utilisant la protéine SAMHD1 humaine entière, a démontré que celle-ci possède une activité nucléase contre l'ADN simple brin et l'ARN simple brin. Son action sur ces deux substrats dépend de la présence de cations métalliques. Cette activité a été caractérisée comme étant une activité exonucléase 3' vers 5' clivant les nucléotides un par un. SAMHD1 semble être capable de dégrader les ARN et ADN simples brins linéaires indépendamment de leur taille, y compris ceux codant pour les protéines Gag et Tat du VIH-1. En présence d'hétéroduplexes ARN-ADN, SAMHD1 clive préférentiellement l'ARN (Beloglazova et al., 2013).

L'activité exonucléase de SAMHD1 repose sur le domaine HD mais est fortement limitée en l'absence du domaine SAM ce qui, dans les conditions de cette étude et contrairement à ce qui avait été montré précédemment (Powell et al., 2011), est également le cas de l'activité dNTPase sur le dGTP (Beloglazova et al., 2013).

L'étude des interactions entre SAMHD1 et les acides nucléiques simples brins montre que celles-ci sont d'autant plus fortes qu'ils contiennent des structures secondaires complexes. Contrairement à l'étude précédemment citée (Goncalves et al., 2012), les auteurs observent cette fois-ci un rôle prépondérant du domaine SAM dans l'interaction avec les acides nucléiques par comparaison avec le domaine HD (Beloglazova et al., 2013).

Une autre étude a caractérisé l'activité nucléase de SAMHD1 *in vitro* et ses auteurs ont observé une activité nucléase uniquement sur les ARN simples brins mais pas sur les ARN doubles brins, les ADN simples brins, les ADN doubles brins ou les hétéroduplexes (Ryoo et al., 2014).

D'importantes différences sont observées entre ces études réalisées *in vitro*. Certaines n'identifient pas d'activité nucléase de SAMHD1 (Goldstone et al., 2011; Goncalves et al., 2012) alors que d'autres observent une activité nucléase et la caractérisent très finement (Beloglazova et al., 2013; Ryoo et al., 2014). L'importance des domaines SAM et HD dans l'activité dNTPase de SAMHD1 et dans sa capacité à interagir avec les acides nucléiques varie également selon les études. Cela peut s'expliquer par un certain nombre de différences dans les paramètres expérimentaux (travail sur des fragments de SAMHD1 ou sur la protéine entière, utilisation d'ions différents à des concentrations variables, origines différentes des acides nucléiques substrats...) et la meilleure façon de répondre définitivement à ces contradictions serait d'étudier l'activité nucléase de SAMHD1 *in vivo*.

Une seule étude a pour l'instant mis en évidence une activité nucléase de SAMHD1 *in vivo*, dans des lignées cellulaires mais également dans des macrophages et des lymphocytes CD4+ quiescents. Cette activité conduit à la dégradation du génome du VIH-1 dès les phases précoces de la transcription inverse, mais le substrat exact et le type d'activité nucléase n'ont pas été caractérisés *in vivo* (Ryoo et al., 2014).

c/ Articulation de ces deux activités

Une seule étude portant sur les deux activités de SAMHD1 a été publiée à ce jour. Celle-ci avait pour objectif de discriminer les deux activités catalytiques de SAMHD1 et apporte donc peu d'information sur la façon dont elles s'articulent (Ryoo et al., 2014).

Il est intéressant de noter que les deux études ayant mis en évidence une activité nucléase de SAMHD1 ont montré que celle-ci est inhibée par l'ajout de dGTP (Beloglazova et al., 2013; Ryoo et al., 2014) et que l'une d'entre elle montre que le dGTP inhibe également l'interaction de SAMHD1 avec les acides nucléiques (Beloglazova et al., 2013). En revanche la présence d'ARN ou d'ADN simple brin ne modifie pas l'activité dNTPase de SAMHD1 (Amie et al., 2013).

En partant de ces données, on pourrait concevoir un modèle selon lequel le GTP et les dNTP, en promouvant la formation de multimères SAMHD1, inhibent l'activité nucléase de cette enzyme (Ryoo et al., 2014) tandis que les acides nucléiques simple brin, n'induisant pas la dislocation des multimères SAMHD1, n'inhibent pas l'activité dNTPase de cette enzyme.

Il faut cependant souligner qu'une étude a montré que des mutants de SAMHD1 ayant perdu la capacité d'interagir avec les acides nucléiques perdent également leur aptitude à former des multimères. Si la forme majoritaire de SAMHD1 observée in vivo par ces chercheurs est le dimère, ils observent également des tétramères et des octamères qui ne se forment qu'en présence d'acides nucléiques (Tüngler et al., 2013). Ces données montrent que les acides nucléiques jouent un rôle dans la multimérisation de SAMHD1 et soulignent le fait que le modèle ci-dessus n'explique pas l'effet des acides nucléiques sur SAMHD1 et donne peu d'information sur la régulation de l'activité nucléase de cette protéine.

La structure de SAMHD1 étant désormais connue et des mutants permettant de séparer l'activité nucléase et l'activité dNTPase ayant été caractérisés, l'articulation de ces deux activités pourra être mieux comprise.

3/ Une régulation encore peu connue

Si les mécanismes moléculaires de la régulation de SAMHD1 par la protéine virale Vpx sont largement étudiés et de mieux en mieux caractérisés, peu de données sont disponibles sur la régulation de SAMHD1 en dehors du contexte de l'infection par le VIH.

a/ Régulation transcriptionnelle et post-transcriptionnelle

En ce qui concerne l'expression du gène correspondant, il a été montré que celle-ci peut-être inhibée par méthylation du promoteur (De Silva et al., 2013a; De Silva et al., 2013b). Ainsi, le promoteur de SAMHD1 n'est pas méthylé dans certaines lignées cellulaires où la protéine et l'ARNm correspondants sont exprimés (comme les THP-1) ni dans les lymphocytes CD4+ mais il est méthylé dans certaines lignées cellulaires où SAMHD1 n'est pas exprimé (comme les Jurkat ou les SupT1) (De Silva et al., 2013b). La méthylation du promoteur de SAMHD1 est à l'origine de l'inhibition de l'expression de la protéine correspondante chez les patients souffrant de certains lymphomes T cutanés (De Silva et al., 2013a).

De façon intéressante, la méthylation du promoteur SAMHD1 conduit à l'absence d'expression de la protéine correspondante mais, dans des lignées tout comme chez des patients, il existe des cas où la protéine ne s'exprime pas alors que le promoteur n'est pas méthylé. Cette observation suggère qu'il existe d'autres mécanismes de régulation de l'expression de SAMHD1.

Il est connu que l'expression de SAMHD1 est induite par l'IFN (Berger et al., 2011a; Chen et al., 2014). D'autres cytokines peuvent également induire l'expression de cette protéine. Ainsi, le traitement de macrophages par une combinaison d'interleukine (IL)-12 et d'IL-18 induit la surexpression des ARNm et des protéines SAMHD1 ce qui rend ces cellules résistantes à l'infection par le VIH-1 (Pauls et al., 2013). Le mécanisme mis en jeu n'est pas caractérisé, mais différentes cytokines pro-inflammatoires sont capables d'induire l'expression de SAMHD1.

L'expression de SAMHD1 est également régulée au niveau post-transcriptionnel. En effet, il a été montré qu'un microARN (miR), le miR181, induit la dégradation des ARNm SAMHD1 dans différentes lignées cellulaires. Le fait que le traitement de lignées exprimant SAMHD1, comme les THP-1, avec un inhibiteur du miR181 augmente l'expression de SAMHD1 indique que l'expression de cette protéine n'est pas régulée par ce miR que dans les lignées où son expression est réduite (Jin et al., 2014).

b/ Régulation post-traductionnelle

La régulation post-traductionnelle de SAMHD1 est essentiellement médiée par des phosphorylations. Différents sites de phosphorylation potentiels ont été identifiés, parmi lesquels la sérine 6, la thréonine 21, la sérine 33 et la thréonine 592 (Welbourn et al., 2013; White et al., 2013b) (numérotation des acides aminés par rapport à la protéine humaine). Pour les trois premiers sites, il a été montré que des protéines mutantes phospho-mimétiques ou ne pouvant plus être phosphorylées ont les mêmes propriétés que SAMHD1 sauvage en terme de localisation, de sensibilité à Vpx, d'activité dNTPase et de restriction du VIH-1 (Welbourn et al., 2013). La phosphorylation de la T592 a été plus étudiée car elle abolit la restriction du VIH-1 par SAMHD1 (Cribier et al., 2013; Welbourn et al., 2013; White et al., 2013b).

Cette T592 est localisée au sein d'un motif susceptible d'être reconnu par deux types de kinase : les kinases dépendantes des cyclines (CDK) et les protéines kinases activées par les agents mitogènes (MAPK) (Welbourn et al., 2013). La phosphorylation de la T592 de SAMHD1 ne modifie ni sa sensibilité à Vpx, ni sa localisation, ni son activité dNTPase, ni sa capacité d'oligomérisation, ni son interaction avec les ARN double brin (Welbourn et al., 2013; White et al., 2013b).

SAMHD1 est phosphorylé sur la T592 dans les cellules sujettes à des divisions cellulaires mais pas dans les cellules différenciées ou quiescentes, qu'il s'agisse de lignées ou de cellules primaires (Cribier et al., 2013; Pauls et al., 2014a; White et al., 2013b). Le traitement de cellules THP-1, de lymphocytes CD4+ activés, de macrophages ou de cellules dendritiques avec de l'IFN α induit une diminution de la phosphorylation sur la T592. A l'inverse, l'activation de PBMC ou de lymphocytes CD4+ par un traitement CD3/CD28/IL2 entraîne une augmentation de la phosphorylation sur la T592 (Cribier et al., 2013).

La principale controverse concernant la phosphorylation de la T592 de SAMHD1 concerne l'identification de la kinase qui la catalyse. Initialement, il a été montré que SAMHD1 interagit avec le complexe cycline A2/CDK1 dans les cellules en prolifération et que ce complexe peut phosphoryler la T592 in vitro (Cribier et al., 2013). Cette capacité de CDK1 à phosphoryler SAMHD1 in vitro a également été observée en association avec la cycline B (White et al., 2013b).

L'interaction entre SAMHD1 et CDK2 a ensuite été caractérisée (Pauls et al., 2014a; St Gelais et al., 2014). Une première étude, réalisée dans des lignées cellulaires en utilisant des inhibiteurs et des siRNA (ARN interférents) spécifiques de CDK1 et CDK2, a permis de montrer que ces deux CDK catalysent la phosphorylation de la T592 de SAMHD1 (St Gelais et al., 2014). Une seconde étude a formellement démontré que CDK2 peut phosphoryler SAMHD1 in vitro, essentiellement mais pas uniquement au niveau de la T592. Cette étude s'est intéressée à la phosphorylation de SAMHD1 dans les macrophages et montre que, dans ces cellules, des siRNA ciblant CDK2 diminuent la phosphorylation de la T592 contrairement aux siRNA ciblant CDK1 (Pauls et al., 2014a).

Les kinases CDK1 et CDK2 possèdent toutes deux la capacité de phosphoryler SAMHD1 in vitro et il semble que leurs importances respectives in vivo varient selon les types cellulaires considérés. Dans tous les cas, le rôle des complexes cycline/CDK dans la régulation post-traductionnelle de SAMHD1 est avéré et montre que cette régulation est intimement liée à celle du cycle cellulaire.

Pour finir, un mécanisme de régulation de SAMHD1 par adressage au protéasome indépendamment de la protéine Vpx a été identifié. La cycline L2 est une protéine cellulaire impliquée dans la régulation du cycle cellulaire et dans l'épissage. Elle interagit avec Dcaf1 et SAMHD1 et est ainsi capable de recruter SAMHD1 au niveau de l'ubiquitine-ligase CRL4 et d'induire sa dégradation par le protéasome. Ce mécanisme semble être important lors de l'infection des macrophages par le VIH-1 (Kyei et al., 2015).

4/ Fonction(s) et mécanismes d'action

a/ Mécanisme antiviral en cours d'élucidation

Suite à la découverte de SAMHD1 et à son identification en tant que dNTPase, plusieurs études ont défendu un modèle selon lequel SAMHD1 restreint le VIH en diminuant la quantité de dNTP disponibles dans les cellules. Cette limitation des dNTP nécessaire à la synthèse de l'ADN viral bloquerait la transcription inverse. Ce modèle est en adéquation avec les observations montrant que SAMHD1 restreint la réplication du VIH dans les cellules quiescentes ou différenciées et non dans les cellules se divisant car ces dernières ont besoin de dNTP pour la réplication de leur propre génome.

Les études ayant conduit à l'émergence de ce modèle s'appuient sur des résultats comparables. D'une part le traitement de cellules réfractaires à l'infection par le VIH avec la protéine Vpx induit la dégradation de SAMHD1 et conduit à une augmentation de la quantité de dNTP intracellulaires. D'autre part l'ajout de dNTP dans le milieu de culture de cellules dont SAMHD1 restreint l'infection par le VIH rend ces cellules permissives à l'infection (Kim et al., 2012; Lahouassa et al., 2012; St Gelais et al., 2012).

L'idée que la quantité de dNTP peut être modulée pour limiter l'infection par le VIH a également été soulignée par l'étude de p21. Cette protéine limite l'infection des macrophages par le VIH-1. Il a été montré que ce phénotype dépend en partie de SAMHD1, car p21 étant un inhibiteur de CDK1 et de CDK2, sa suppression augmente donc la phosphorylation de SAMHD1 sur la T592 ce qui rend les macrophages permissifs au VIH (Pauls et al., 2014b). Mais ce phénotype dépend également de la ribonucléotide diphosphate réductase 2 (RNR2). En effet, la transcription de cette enzyme impliquée dans la synthèse des dNTP est inhibée par p21, ce qui limite l'infection des macrophages par le VIH-1 (Allouch et al., 2013).

La principale limite de ce modèle a été découverte lors de l'étude de la régulation de SAMHD1 par phosphorylation sur la T592. En effet, si l'ensemble des travaux réalisés sur cette phosphorylation montre qu'elle supprime la capacité de SAMHD1 à restreindre l'infection par le VIH (Cribier et al., 2013; Pauls et al., 2014a; St Gelais et al., 2014; Welbourn et al., 2013; White et al., 2013b), il montre également que cette phosphorylation ne diminue pas l'activité dNTPase de SAMHD1 (Welbourn et al., 2013; White et al., 2013b). Or, s'il est possible que SAMHD1 perde sa capacité à bloquer le VIH en conservant son activité dNTPase, c'est que cette activité dNTPase n'est pas à l'origine de la restriction causée par SAMHD1.

Une étude récente utilisant des mutants de SAMHD1 ne possédant qu'une seule des deux activités catalytiques de cette enzyme a permis de montrer que l'activité dNTPase n'est pas impliquée dans le phénotype de restriction, à l'inverse de l'activité nucléase (Ryoo et al., 2014). Lors de l'infection par le VIH, SAMHD1 semble capable de dégrader directement le génome viral avant l'aboutissement de la transcription inverse.

Ce modèle est cohérent avec l'observation précédemment décrite selon laquelle l'ajout de dNTP exogènes supprime l'activité de restriction de SAMHD1. En effet, il a été montré par plusieurs études indépendantes que les dNTP induisent la multimérisation de SAMHD1 et que ces multimères sont actifs en tant que dNTPases mais pas en tant que nucléases. L'ajout de dNTP exogènes supprime donc la capacité de restriction de SAMHD1 en inhibant son activité nucléase et non pas en compensant son activité dNTPase.

Ce modèle n'est pas non plus en contradiction avec le fait que la dégradation de SAMHD1 par Vpx conduise à une augmentation de la concentration des dNTP intracellulaires puisqu'il n'exclut pas qu'une fraction de la protéine SAMHD1 soit active en tant que dNTPase tandis qu'une autre fraction, active en tant que nucléase, restreigne l'infection par le VIH.

En revanche, il a été observé que la capacité de SAMHD1 à former des tétramères est nécessaire pour que cette protéine puisse restreindre l'infection par le VIH-1 (Zhu et al., 2013). Or à l'heure actuelle la forme tétramérique de SAMHD1 est considérée comme active en tant que dNTPase et inactive en tant que nucléase.

En résumé, la compréhension du mécanisme exact à l'origine de la restriction de différents virus par SAMHD1 ne sera vraisemblablement possible que quand l'articulation des activités catalytiques de cette enzyme aura été finement caractérisée. A l'heure actuelle, certains résultats expérimentaux semblent être contradictoires et, si l'ensemble des données suggère un rôle majoritaire de l'activité nucléase et un rôle minoritaire de l'activité dNTPase dans la restriction, il reste difficile de concevoir un modèle qui explique l'ensemble des observations.

Dans la plupart des études concernant la phosphorylation de SAMHD1, une corrélation parfaite a été observée entre la phosphorylation sur la T592 et la perte de la capacité à restreindre l'infection par le VIH. Ceci laisse penser que l'activité de restriction de SAMHD1 n'est observée que dans les cellules ne se divisant pas car celles-ci n'expriment pas les protéines phosphorylant la T592 de SAMHD1. Autrement dit que SAMHD1, bien qu'exprimé dans certaines cellules en division, y est inhibé par phosphorylation et n'y restreint donc pas l'infection par le VIH.

Il existe néanmoins une exception puisque, dans une des études, l'expression d'un mutant de SAMHD1 non phosphorylable sur la T592 dans des cellules 293T n'a pas permis de rendre ces dernières réfractaires à l'infection par le VIH (St Gelais et al., 2014). Cette observation prouve que la capacité de SAMHD1 à restreindre l'infection par le VIH ne dépend pas uniquement de sa phosphorylation sur la T592. Il serait particulièrement intéressant d'étudier l'effet de la phosphorylation des autres sites phosphorylables et d'éventuelles autres modifications post-traductionnelles de cette protéine sur cette propriété.

b/ Fonction(s) cellulaire(s) peu voire non connues

Lors de l'identification de SAMHD1 comme un facteur de restriction du VIH, peu de choses étaient connues sur la fonction cellulaire de cette protéine en dehors d'un contexte infectieux. On savait cependant que des mutations dans le gène SAMHD1 pouvaient être à l'origine d'une maladie auto-immune, le syndrome Aicardi-Goutière (AGS) (Rice et al., 2009). Cette maladie touchant essentiellement les nourrissons est caractérisée par un fort taux d'IFN α , notamment dans le liquide céphalorachidien, et une hyper inflammation qui rappellent les symptômes des infections virales. L'implication de mutations de SAMHD1 dans cette maladie ainsi que dans d'autres pathologies inflammatoires (Dale et al., 2010; Ravenscroft et al., 2011) suggère que la protéine correspondante pourrait être un régulateur négatif de la réponse immunitaire innée.

Des mutations dans 8 gènes différents ont été identifiées chez les patients AGS. TREX1, muté chez 23% des patients, code pour une exonucléase dégradant les ADN double et simple brin. La RNaseH2, dont les gènes codant pour les sous-unités A, B et C sont mutés respectivement chez 5%, 36% et 12% des patients, est une endonucléase nucléaire dégradant l'ARN dans les hétéroduplexes. ADAR, muté chez 7% des patients, code pour une adénosine déaminase éditant les ARN double brin tandis que MDA5, ou IFIH1, muté chez 3% des patients, code pour un senseur d'ARN double brin capable de promouvoir la réponse immunitaire innée. Enfin, les mutations dans SAMHD1 sont à l'origine de 13% des cas d'AGS actuellement répertoriés (Crow et al., 2015). L'ensemble des gènes impliqués dans l'AGS jouent un rôle dans le métabolisme des acides nucléiques, ce qui laissait penser que cela devait être le cas pour SAMHD1 avant que ses activités dNTPase et nucléase soient caractérisées.

A l'heure actuelle on suppose que SAMHD1 pourrait limiter l'expression des éléments mobiles du génome, que cela soit via son activité dNTPase ou via son activité nucléase, et ainsi empêcher l'accumulation d'acides nucléiques anormaux dans le cytoplasme. En l'absence de SAMHD1, ces acides nucléiques seraient détectés par des senseurs qui entraîneraient une réponse immunitaire en l'absence de pathogène. Ce modèle est en accord avec l'ensemble des données expérimentales concernant SAMHD1 mais n'a pas été formellement démontré.

Les patients souffrant d'un AGS causé par une mutation dans le gène SAMHD1 développent une forme particulièrement sévère de la maladie, associée à des problèmes vasculaires dans le cerveau pouvant causer des hémorragies cérébrales et autres accidents vasculaires cérébraux (Ramesh et al., 2010; Thiele et al., 2010; Xin et al., 2011). Ces données suggèrent qu'en plus de son rôle dans la régulation de l'immunité innée et dans le métabolisme des acides nucléiques, SAMHD1 pourrait avoir un rôle dans la mise en place ou la régulation du fonctionnement des vaisseaux sanguins.

Deux modèles de souris n'exprimant pas le gène SAMHD1 ont été mis en place. Malheureusement, ils ne permettent pas d'étudier le mécanisme d'établissement de l'AGS en l'absence de SAMHD1. En effet, dans les deux cas, les souris ne développent pas de pathologie auto-immune comparable à l'AGS, même si une augmentation d'IFN et des ISG a été mise en évidence dans certains tissus (Behrendt et al., 2013; Rehwinkel et al., 2013). Cela suggère que les fonctions cellulaires de SAMHD1 pourraient différer chez l'homme et chez la souris.

Un modèle de poisson zèbre n'exprimant pas SAMHD1 a également été mis en place et celui-ci a un phénotype plus proche de l'AGS que celui des souris. En effet, les poissons surexpriment des gènes de la réponse immunitaire innée, dont les IFN de type I et développent des hémorragies cérébrales (Kasher et al., 2015).

L'utilisation combinée de ces différents modèles devrait permettre de mieux comprendre les liens concrets entre SAMHD1, la réponse immunitaire innée et l'irrigation sanguine du cerveau.

A l'échelle cellulaire, des données de plus en plus précises sont disponibles concernant la fonction de SAMHD1 au niveau du cycle cellulaire. SAMHD1 est régulée au niveau post-traductionnel par phosphorylation mais également en terme de niveau d'expression au cours du cycle cellulaire (Kretschmer et al., 2014). Cette régulation est inversement coordonnée avec celle de la RNR, ce qui permet de finement contrôler la quantité de dNTP disponible au cours de la phase S pour la réplication du génome cellulaire et hors de cette phase S (Franzolin et al., 2013).

Si le cycle cellulaire a un impact sur la régulation de SAMHD1, il a également été montré que la dérégulation de SAMHD1 a un impact sur le cycle cellulaire. En effet, la diminution de l'expression de SAMHD1 ralentit la prolifération cellulaire et conduit à une accumulation de cellules en phase G1, ce qui suggère un défaut dans la transition G1/S en l'absence de SAMHD1 (Franzolin et al., 2013; Kretschmer et al., 2014). En parallèle, la quantité disponible des quatre dNTP augmente et la proportion relative de ces dNTP est modifiée, avec notamment une forte augmentation de la proportion de dGTP (au moins deux fois) (Franzolin et al., 2013). Ces perturbations du cycle cellulaire en l'absence de SAMHD1 sont couplées à l'induction d'inhibiteurs du cycle cellulaire comme la protéine p21 (Kretschmer et al., 2014).

En plus des liens entre SAMHD1, l'immunité innée et la progression du cycle cellulaire, différentes données suggèrent que SAMHD1 pourrait avoir un rôle plus large dans l'homéostasie cellulaire. En effet, il a été observé que la mutation du gène SAMHD1 chez certains patients conduit à des délétions dans l'ADN mitochondrial (Leshinsky-Silver et al., 2011). Ce génome mitochondrial est répliqué indépendamment du cycle cellulaire et l'absence de SAMHD1 semble perturber ce processus.

Par ailleurs, dans les fibroblastes de patients AGS n'exprimant pas SAMHD1, des perturbations autres que celles concernant le cycle cellulaire ont pu être observées. Notamment une activité accrue des lysosomes, la formation de foci d'hétérochromatine dans le noyau et une augmentation globale des dommages à l'ADN, qui sont autant de marqueurs de sénescence prématurée (Kretschmer et al., 2014). La perte de SAMHD1 semble donc affecter différents aspects du fonctionnement cellulaire.

Une étude menée sur des poissons zèbres a permis de montrer que le traitement d'embryons avec de l'atrazine, un perturbateur endocrinien avéré et oncogène potentiel, induit la surexpression de SAMHD1 (Weber et al., 2013). Ainsi, l'absence de SAMHD1 entraîne de nombreuses perturbations du fonctionnement cellulaire mais la surexpression de SAMHD1 pourrait elle-même constituer un mécanisme de réponse à des perturbations du fonctionnement de l'organisme.

En résumé, si aucune fonction cellulaire associée à un mécanisme d'action n'a clairement été définie pour SAMHD1, les données disponibles suggèrent que cette protéine possède plusieurs fonctions importantes. Tout d'abord SAMHD1 est vraisemblablement impliquée dans la régulation de la réponse immunitaire innée et pourrait moduler l'irrigation sanguine du cerveau. Par ailleurs SAMHD1 module le cycle cellulaire, notamment via son activité dNTPase et est elle-même régulée pendant le cycle cellulaire. Enfin, SAMHD1 semble jouer un rôle plus large dans le contrôle de l'homéostasie cellulaire, influençant l'intégrité du génome mitochondrial, l'organisation de la chromatine et l'activité lysosomale.

Problématique

Depuis leur découverte, les virus ont constitué un outil privilégié pour mieux comprendre le monde du vivant. Ils sont à l'origine de la caractérisation de nombreux mécanismes fondamentaux et peuvent désormais être utilisés comme outils expérimentaux ou thérapeutiques. Les rétrovirus et notamment le VIH ont contribué à attirer l'attention des chercheurs sur plusieurs protéines cellulaires dont le fonctionnement reste encore mal compris. C'est notamment le cas des facteurs de restriction.

Ces protéines ont co-évolué avec les virus qu'elles restreignent mais il est clair que leurs fonctions ne sont pas uniquement antivirales. Au-delà de leur histoire évolutive, le fait que certains facteurs de restriction soient constitutivement exprimés dans la plupart des cellules humaines en l'absence d'infection souligne leur implication dans des fonctions complémentaires.

Lors de l'identification de SAMHD1 comme étant un facteur de restriction du VIH-1, le gène correspondant était déjà connu comme impliqué dans une maladie auto-immune : le syndrome Aicardi-Goutière. Or l'ensemble des gènes pouvant être à l'origine de ce syndrome codent des protéines impliquées dans le métabolisme des acides nucléiques et la réponse immunitaire innée. Cette observation, couplée au fait que SAMHD1 restreint le VIH au niveau de l'étape de transcription inverse, suggérait déjà un rôle de SAMHD1 dans le métabolisme des acides nucléiques.

Deux activités catalytiques ont depuis été caractérisées pour SAMHD1, effectivement en lien avec les acides nucléiques : cette protéine est une dNTP triphosphohydrolase et une nucléase. Cependant l'essentiel des travaux réalisés sur SAMHD1 portent sur son activité antivirale et, au début de ma thèse, très peu de données étaient disponibles sur les possibles fonctions cellulaires de cette protéine.

Considérant le fait que SAMHD1 joue un rôle dans le métabolisme des acides nucléiques et particulièrement des dNTP, j'ai voulu savoir si cette protéine modulait les processus cellulaires nécessitant des dNTP.

En effet, les dNTP interviennent dans des processus aussi fondamentaux que la réplication du génome et la réparation des dommages à l'ADN. J'ai ainsi pu montrer que la surexpression de SAMHD1 ralentit la prolifération cellulaire et favorise la mort cellulaire en cas de rupture double brin de l'ADN. Les mécanismes correspondants restent à caractériser, mais mes résultats montrent que SAMHD1 est effectivement impliquée dans des processus cellulaires essentiels.

Or la dérégulation de ces processus peut être à l'origine de cancers. De fait, la capacité de prolifération illimitée et la survie en cas de lésions à l'ADN sont deux marqueurs des cellules cancéreuses. Je me suis donc intéressée à l'expression de SAMHD1 dans différents cancers primaires.

J'ai observé que, par différents mécanismes, l'expression de SAMHD1 est réduite dans un certain nombre de cancers primaires et notamment chez 80% des patients souffrant de leucémies lymphoïdes chroniques. Cette fréquence importante et les résultats préalablement obtenus *in vitro* suggèrent que la perte de SAMHD1 n'est pas une conséquence aléatoire de l'instabilité génétique des cellules cancéreuses mais pourrait au contraire être un élément moteur de l'évolution vers le cancer.

L'ensemble de ces résultats souligne le fait que SAMHD1 n'est pas seulement un facteur de restriction actif contre différents virus. C'est avant tout une protéine essentielle à la régulation de processus cellulaires fondamentaux dont l'altération innée ou acquise peut être liée au développement de pathologies aussi graves que l'AGS et le cancer. Les mécanismes sous-jacents et le lien de SAMHD1 avec la réponse immunitaire innée restent à élucider.

Résultats

Éléments de contextualisation

Les résultats présentés dans ce manuscrit correspondent à des travaux de recherche débutés en 2012 et qui ont abouti à une publication dans la revue *Blood* fin 2013. D'autres projets ont été menés en parallèle mais ceux-ci étant encore en cours au sein du laboratoire, nous avons fait le choix de ne pas les détailler ici.

Lorsque le projet présenté ici a été lancé, très peu de choses étaient connues sur le fonctionnement et la régulation de SAMHD1 en l'absence d'infection par le VIH. L'activité dNTPase de cette protéine était caractérisée et on connaissait le rôle des mutations dans le gène correspondant dans le développement de l'AGS. L'existence d'une activité nucléase de SAMHD1 *in vitro*, l'identification de la régulation par phosphorylation sur la T592 et les liens entre SAMHD1 et le cycle cellulaire sont autant de résultats qui ont été publiés pendant le développement de mon projet de thèse.

Enfin, ces travaux ont été publiés en collaboration avec l'équipe d'Anna Schuh. L'article correspondant a été rédigé en co-premier et co-dernier auteur. Ainsi, seule une partie des données présentées dans l'article est le fruit de mon travail. Les légendes correspondantes ont été surlignées en jaune dans le manuscrit ci-après. Afin de clarifier le message de l'article, celui-ci a été rédigé dans une logique intégrant les résultats des deux équipes. C'est pourquoi, après la présentation du manuscrit, je tiens à replacer les résultats présentés dans la logique qui a initialement été la mienne.

Article publié

Regular Article

LYMPHOID NEOPLASIA

***SAMHD1* is mutated recurrently in chronic lymphocytic leukemia and is involved in response to DNA damage**

Ruth Clifford,¹ Tania Louis,² Pauline Robbe,¹ Sam Ackroyd,³ Adam Burns,¹ Adele T. Timbs,¹ Glen Wright Colopy,¹ Helene Dreau,¹ Francois Sigaux,⁴ Jean Gabriel Judde,⁵ Margalida Rotger,⁶ Amalio Telenti,⁶ Yea-Lih Lin,² Philippe Pasero,² Jonathan Maelfait,⁷ Michalis Titsias,⁸ Dena R. Cohen,⁹ Shirley J. Henderson,¹ Mark T. Ross,¹⁰ David Bentley,¹⁰ Peter Hillmen,¹¹ Andrew Pettitt,¹² Jan Rehwinkel,⁷ Samantha J. L. Knight,¹³ Jenny C. Taylor,¹³ Yanick J. Crow,¹⁴ Moncef Benkirane,² and Anna Schuh¹

¹Oxford National Institute for Health Research Biomedical Research Centre/Molecular Diagnostic Centre, University of Oxford, Oxford, United Kingdom;

²Institut de Génétique Humaine, Centre National de la Recherche Scientifique-Unités Propres de Recherche 1142, Montpellier, France; ³Clinical

Haematology, Bradford Teaching Hospitals, National Health Service Foundation Trust, Bradford, United Kingdom; ⁴Laboratory of Molecular Hematology,

Institut National de la Santé et de la Recherche Médicale U462, Centre Hayem, Hôpital Saint Louis, Paris, France; ⁵Xentech SAS, Genopole, Evry, France;

⁶Institute of Microbiology, University Hospital Center and University of Lausanne, Lausanne, Switzerland; ⁷Medical Research Council Human Immunology

Unit, Radcliffe Department of Medicine, Medical Research Council Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, United Kingdom;

⁸Athens University of Economics and Business, Athens, Greece; ⁹Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, United Kingdom;

¹⁰Illumina Cambridge Ltd., Saffron Walden, United Kingdom; ¹¹St. James' Institute of Oncology, Leeds, United Kingdom; ¹²Department of Molecular and

Clinical Cancer Medicine, University of Liverpool, Liverpool, United Kingdom; ¹³Oxford National Institute for Health Research Biomedical Research Centre

and Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; and ¹⁴Genetic Medicine, Institute of Human Development,

Faculty of Medical and Human Sciences, University of Manchester, St. Mary's Hospital, Central Manchester University Hospitals National Health Service

Foundation Trust, Manchester Academic Health Science Centre, Manchester, United Kingdom

Key Points

- Acquired pathogenic mutations in *SAMHD1* are found in up to 11% of relapsed/refractory patients with CLL.
- *SAMHD1* is mobilized to sites of DNA damage.

***SAMHD1* is a deoxynucleoside triphosphate triphosphohydrolase and a nuclease that restricts HIV-1 in noncycling cells. Germ-line mutations in *SAMHD1* have been described in patients with Aicardi-Goutières syndrome (AGS), a congenital autoimmune disease. In a previous longitudinal whole genome sequencing study of chronic lymphocytic leukemia (CLL), we revealed a *SAMHD1* mutation as a potential founding event. Here, we describe an AGS patient carrying a pathogenic germ-line *SAMHD1* mutation who developed CLL at 24 years of age. Using clinical trial samples, we show that acquired *SAMHD1* mutations are associated with high variant allele frequency and reduced *SAMHD1* expression and occur in 11% of relapsed/refractory CLL patients. We provide evidence that *SAMHD1* regulates cell**

proliferation and survival and engages in specific protein interactions in response to DNA damage. We propose that *SAMHD1* may have a function in DNA repair and that the presence of *SAMHD1* mutations in CLL promotes leukemia development. (*Blood*. 2014;123(7):1021-1031)

Introduction

Chronic lymphocytic leukemia (CLL) is a disease of the elderly, demonstrating both clinical and biological heterogeneity. Complications include the development of autoimmunity and secondary immunodeficiency. Recurrent copy number aberrations of prognostic significance in CLL have been identified by fluorescence in situ hybridization,¹ including deletions of 17p13.1 involving *TP53* and deletions of 11q22.3 involving *ATM*. Recent data from whole genome and exome sequencing (WGS and WES, respectively) have confirmed the heterogeneous nature of CLL, with all recurrent mutations described occurring at low frequency.²⁻⁵

Previously, we used WGS to perform longitudinal studies of CLL patients undergoing treatment and described likely pathogenic variants in candidate founder genes that might initiate and/or maintain leukemia.⁶ These mutations were somatic and defined by their presence

at all 5 time points in all leukemia cells assayed.⁶ Of particular interest was an acquired uniparental isodisomy event with a homozygous nonsynonymous mutation in *SAMHD1* (supplemental Figure 1 on the *Blood* Web site), a gene involved in the regulation of innate immune response and autoimmune disease.

Germ-line mutations in *SAMHD1* have been described in 17% of patients diagnosed with a congenital autoimmune disease called Aicardi-Goutières syndrome (AGS).^{7,8} AGS is an autosomal recessive disorder that is characterized by an autoimmune encephalopathy and chronic activation of the immune system. It occurs mainly in patients of consanguineous background. Most patients with AGS succumb to the disease in early childhood.

More recently, *SAMHD1* has been identified as an HIV-1 restriction factor operating in nondividing cells including dendritic

Submitted March 29, 2013; accepted November 17, 2013. Prepublished online as *Blood* First Edition paper, December 12, 2013; DOI 10.1182/blood-2013-04-490847.

R.C., T.L., M.B., and A.S. contributed equally to this work.

The online version of this article contains a data supplement.

There is an Inside *Blood* commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2014 by The American Society of Hematology

cells, macrophages, and resting CD4⁺ T cells.⁹⁻¹² SAMHD1 is a deoxyribonucleotide triphosphate (dNTP) triphosphohydrolase and a nuclease^{13,14} that regulates the intracellular dNTP pool.^{15,16} It is expressed widely across most tissue types including in the hemopoietic system and normal B cells.^{17,18} In the case of HIV, SAMHD1 is responsible for maintaining the cellular dNTP pool at a level that is inadequate for replication of HIV-1 and other viruses,¹⁹⁻²¹ thereby preventing viral propagation.

Importantly, there is increasing evidence that an imbalance in the cellular dNTP pool is also associated with DNA replication stress and an alteration in replication fork velocity. Perturbations of the replication machinery in turn have been shown to lead to increased mutagenesis, genomic instability, and cancer development.^{22,23}

Taken together, (1) the potential role of a *SAMHD1* mutation as a founder event in CLL, (2) the established function of SAMHD1 in maintenance of the intracellular dNTP pools, and (3) the involvement of SAMHD1 in the innate immune response and AGS encouraged us to further explore the role of *SAMHD1* in CLL pathogenesis.

To this, we present a detailed molecular characterization of clinical trial samples from CLL patients in the United Kingdom together with functional experiments interrogating the role of SAMHD1 in the presence of DNA damage.

Materials and methods

Samples

Informed consent from CLL patients was obtained in line with the Declaration of Helsinki. Ethics approval for this project is covered under the Haematology Collection Protocol, HTA License Number 12217, Oxfordshire C REC: 09/H0606/5. We initially chose CLL patients from our institution (Oxford University Hospitals National Health Service [NHS] Trust) to evaluate whether *SAMHD1* mutations were recurrent in CLL (cohort 1 = 100 samples). Overall, the patient cohort was that of a tertiary referral center with a high proportion of unmutated *IGVH* and treatment refractory cases. Chemotherapy refractoriness was defined by failure to respond to a purine analog or relapsing within 6 months of treatment.

To evaluate the frequency of *SAMHD1* mutations in a pretreatment cohort, 200 samples from 2 United Kingdom (UK) National Cancer Research Network (NCRN) clinical trials were sequenced (cohort 2: Does the ADDition of Mitoxantrone Improve REsponse (AdMiRE) trial, n = 108; Attenuated dose Rituximab with ChemoTherapy In CLL (ARCTIC) trial, n = 92). A further 63 samples from 2 UK NCRN relapsed and refractory trials (cohort 3: CLL 201, n = 37; CLL 202, n = 26) were also evaluated for *SAMHD1* mutations. All clinical trial samples were sequenced using our in-house next-generation sequencing (NGS) targeted panel (supplemental Table 1)

Peripheral blood mononuclear cells (PBMCs) were isolated from CLL blood samples by ficoll gradient centrifugation. DNA was extracted from peripheral blood CLL cells using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol.

Samples from cohorts 1, 2, and 3 were tested using genome-wide single nucleotide polymorphism (SNP) arrays and NGS, as well as Sanger sequencing. The B-allele frequencies (BAFs) from the array data and the variant allele frequencies (VAFs) from the NGS data showed that the contribution of CLL genomes was ≥80% in all samples.

Sanger sequencing

We designed primers to amplify the coding exons of *SAMHD1* (supplemental Table 2). Each primer had an M13 tail: forward, 5'-GTAAACGACGGC CAGT-3'; reverse, 5'-CAGGAAACAGCTATGAC-3'. Purified amplicons were sequenced from both strands using BigDye terminator chemistry and an ABI 3130 DNA sequencer. Mutation Surveyor DNA variant analysis software was used to compare sequences with germ-line reference sequence.

Genome-wide SNP platform hybridizations and analysis

Hybridization to Illumina SNP chips (HumanOmni1-Quad and HumanOmniS-8 platforms) was performed according to manufacturer's protocols found on registration at <http://www.illumina.com/products>. The data were processed using GenomeStudioV2009.2 (Illumina, San Diego, CA) and analyzed using Nexus 6.1 Discovery Edition (BioDiscovery, El Segundo, CA). We excluded from further analyses all losses and gains that had been noted previously in the Database of Genomic Variants (<http://dgv.tcag.ca/dgv/app/home>) and all copy neutral loss of heterozygosity (cnLOH) <2 Mb.

To determine the genomic complexity of a sample, we performed 2 calculations. (1) The sum of the length of all losses and gains (>20 kb) and cnLOH (>2 Mb). A total length of >100 Mb was used to define genomic complexity. (2) The sum of the number of losses and gains (>20 kb) and cnLOH (>2 Mb). A total of ≥4 was used as the second criterion toward defining genomic complexity. A sample was defined as having genomic complexity only if both criteria were fulfilled.

We defined mosaic cnLOH as the occurrence of a region of allelic imbalance in the absence of a copy number change. Allelic imbalance was detected using the BAF data and defined as present when the heterozygous (AB) allele frequency deviated from the expected 0.5 value but did not reach full AA or BB homozygosity (at 0 and 1 BAF, respectively). We were guided by Nexus Discovery Edition v6.1 calls and reported only changes that were visible on manual inspection of the data plots.

Targeted NGS

A TruSeq Custom Amplicon panel (TSCA; Illumina Cambridge Ltd., Saffron Walden, UK) was designed, targeting mutational hotspots in 11 genes involved in the pathogenesis of CLL (supplemental Table 2). Dual-barcoded TSCA libraries were created from 250 ng of DNA, before undergoing 2 × 150-bp paired-end sequencing on the MiSeq platform (Illumina Cambridge Ltd.). Initial alignment and variant calling analysis was performed with the BaseSpace online analysis tool.²⁴ The sequence data were also analyzed using the Stampy²⁵ and Platypus²⁶ tools to screen for large insertions/deletions. High confidence functional sequence variants were identified by removing any that failed the BaseSpace filter, had a quality score <60, had no predicted functional consequence (intergenic, synonymous), or were present in dbSNP (build 135). All variants with dbSNP identifiers were checked in dbSNP for clinical association.

Cell purification

PBMCs were stained and sorted by fluorescence-activated cell sorter Aria (BD-Pharmingen) as described previously.²⁷ Cells were stained with anti-CD19 phycoerythrin (PE) (BD Pharmingen) and anti-CD5 allophycocyanin (APC) (BD Pharmingen). Postsort purity was >95%.

WES

The tumor sample was purified as described above with the CD19⁺, CD5⁺ fraction selected for WES. The sample was quantified using the High Sensitivity Qubit system (Invitrogen), and sample integrity was assessed using 1% E-Gel EX (Invitrogen). Library preparation was performed using the Illumina TruSeq DNA HT Sample Preparation kit and enriched for the exome using the Illumina TruSeq Exome Enrichment kit. Sequencing was performed using a HiSeq2000 as 100-bp paired end reads.

The sequence reads were mapped to the human genome assembly hg19/NCBI37 using Stampy,²⁵ and variants were detected using Platypus.²⁶ High-quality reads scoring 200, with sequence variations found within coding exons and a total coverage of ≥50 reads, were analyzed further. The database for normal variants (dbSNP)²⁸ and the 1000 Genomes database²⁹ were used to exclude common variants. We searched for any published CLL-associated mutations^{2,4-6,30,31} present in the leukemia cells and excluded their presence in the germ line by Sanger sequencing.

RNA extraction and reverse transcription

Total RNA was isolated from sorted cell populations using RNeasy kits according to the manufacturer's protocol (Qiagen, Hilden, Germany) or

purified with Trizol reagent (Invitrogen). Complementary DNA was prepared from 1 to 12 μ L total RNA using SuperScript VILO kits (Invitrogen, Carlsbad, CA) or SuperScript II enzyme (Invitrogen) and oligo-dT primers according to the manufacturer's protocol.

RNA isoforms

Reverse transcription (RT) products were used for the polymerase chain reaction (PCR) reaction using Phusion HSII enzyme (Finnzymes) and the following SAMHD1-specific primers: CGGCGCCGAGGTTCTTGACT (forward) and TTTGTAAACAAGTACTACAGACA (reverse). The PCR products were resolved on a 1% agarose gel containing 1 $\times$ SYBR Safe DNA gel stain (Life Technologies).

Quantitative real-time RT-PCR of purified cells

To generate Figure 4A, absolute quantification of cDNA was performed in duplicate using real-time PCR and a transcript quantification TaqMan assay kit for *SAMHD1* (Hs00210019_m1; Applied Biosystems, Foster City, CA). Amplifications were performed using Qiagen Rotor-Gene 6000. Following normalization using the control gene, *c-abl*, relative quantification was calculated using the Δ Ct method. Quantification of *SAMHD1* mRNA shown in Figure 4B was performed in triplicate using QuantiTect SYBR Green (Qiagen) according to manufacturer's instructions. Further details are in supplemental Methods.

cDNA sequencing

Primers were designed to amplify the complementary DNA of particular exons of *SAMHD1* (supplemental Table 3). Purified amplicons were sequenced as detailed in the Sanger sequencing method to confirm the presence of the mutations found in the corresponding DNA samples.

Cell lines and expression constructs

Adherent and suspension cells were cultured in Dulbecco's modified Eagle medium or RPMI-1640 respectively, supplemented with 10% fetal calf serum, ultraglutamine, and antibiotics. HeLa cells stably expressing Flag-HA-tagged SAMHD1 have been described previously.⁹ HeLa cells stably expressing inducible SAMHD1 wild-type (WT) and mutants were generated by transducing HeLa cells with lentiviral pTRIPZ vector (Thermo Scientific, Waltham, MA), in which the *SAMHD1* open reading frame (ORF) was inserted. Transduced cells were selected with 2 μ g/mL puromycin. *SAMHD1* expression was induced with 2 μ g/mL doxycycline (Clontech, Saint-Germain-en-Laye, France).

Plasmids

The *SAMHD1* WT plasmid used to produce HeLa cell lines has been described previously.⁹ The SAMHD1 HD/AA mutant and F575 Δ Cter mutant have also been described previously.³²

Cell extracts preparation and western blot analysis

For western blotting, cells were lysed for 30 minutes in a buffer containing 0.5% Triton X-100, 150 mM NaCl, 10 mM KCl, 1.5 mM MgCl₂, 0.5 mM EDTA, 10 mM β -mercaptoethanol, and 0.5 mM phenylmethylsulfonyl fluoride. Equal amount of proteins were resolved on sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The proteins were transferred to nitrocellulose membranes (GE Healthcare, Little Chalfont, Buckinghamshire, UK). When specified, protein load was assessed by Ponceau S staining (Thermo Scientific). Membranes were blocked with 5% milk diluted in phosphate-buffered saline/0.1% Tween, incubated with primary antibody for 1 hour at room temperature, and incubated with horseradish peroxidase-conjugated secondary antibody for 1 hour at room temperature. Proteins were visualized with SuperSignal West pico or femto (Thermo Scientific).

Immunofluorescence

HeLa cells constitutively expressing Flag-HA-tagged SAMHD1 were seeded on cover glasses and treated with or without 25 nM of camptothecin (CPT) for 18 hours. Cells were then fixed with 2% paraformaldehyde and permeabilized using 0.1% Triton X-100. Immunodetection of 53BP1 and HA-tagged SAMHD1 was performed overnight at 4°C using specific antibodies. Images were acquired using the Zeiss confocal microscope system LSM780.

Antibodies

Anti-SAMHD1 antibodies used were purchased from Abcam (Cambridge, UK; ab96768 and ab67820) and Bethyl/Euromedex (Schiltigheim, France; A303-690A). Anti-phosphoSAMHD1Thr592 antibody has been described previously.³³ Anti-CyclinA2 antibody was purchased from Cell Signaling, anti-DCAF1 antibody from Bethyl/Euromedex, anti-HA from Covance/Eurogentec, and anti-53BP1 antibody was from Millipore (Billerica, MA). Tubulin and actin antibodies were from Sigma-Aldrich Chemical (St Louis, MO). Mouse (NA931) and rabbit (NA934) secondary antibodies were purchased from GE Healthcare.

Drug treatment and cell proliferation

HeLa WT or expressing SAMHD1 were seeded in 96-well plates at a density of 1000 cells per well. Etoposide, camptothecin, mitomycin C, and methyl methanesulfonate (Sigma-Aldrich Chemical) were applied to the cells 6 hours after seeding in a total volume of 100 μ L. Three days after treatment, the Cell Titer-Glo kit (Promega, Madison, WI) was used to measure cell survival according to the manufacturer's instructions. Results are shown as percentage of survival compared with mock-treated cells.

For the cell proliferation experiment, the cell count was performed at the indicated time using MACSQuant Analyzer (Miltenyi Biotec, Bergisch Gladbach, Germany).

Cell fractionation and glycerol gradient

HeLa stably expressing SAMHD1 were seeded in 15-cm dishes and treated with 5 μ M etoposide for 4 and 18 hours or mock treated. Cells were then harvested, washed with phosphate-buffered saline, and lysed successively in 5 volumes of CSK buffer (10 mM piperazine-N,N'-bis[2-ethanesulfonic acid] [PIPES], pH 7, 100 mM NaCl, 300 mM sucrose, 3 mM MgCl₂, 10 mM β -glycerophosphate, 0.2 mM Na vanadate, and 5 mM Na fluoride) containing 0.5% and 0.1% of Triton to purify nuclear extracts. For density gradient sedimentation, 400 μ g of nuclear extracts was loaded onto a glycerol gradient (15-35% in buffer containing 20 mM Tris, pH 7.5, 0.15 M KCl, 2.5 mM MgCl₂, 0.05% NP40, 0.05% Tween, 1 mM dithiothreitol, and protease inhibitors; Roche) and centrifuged for 4 hours at 55 000 rpm. Two hundred-microliter fractions were collected from the top of the gradient.

Statistical methods

We wished to ascertain whether *SAMHD1* mutations were associated with common CLL-described mutations (Figure 3B). We analyzed genomic data on 18 *SAMHD1* mutated samples (Figure 3B) and 229 WT *SAMHD1* samples (clinical trial samples with whole genome array and NGS data available). The sample proportions of described CLL mutations were calculated for the mutated and WT SAMHD1 patients, and the difference was tested for significance using the 2-sided pooled hypothesis z-test, which provides a *P* value. Unadjusted *P* values can be compared with an adjusted α level of 0.05/7 = 0.0071 using a conservative Bonferroni adjustment to reflect that 7 comparisons were made.

We compared the clinical outcome for the mutated and WT SAMHD1 patients on samples with clinical outcome available (AdMIRE Trial).³⁴ The 2 populations were compared with respect to the complete remission (CR) and minimal residual disease (MRD), using for both comparisons the standard 2-sided pooled z-test for the difference in proportions.

For patients with both overall survival data and *SAMHD1* mutation status available (CLL 201 and CLL 202 clinical trials, n = 65), a survival

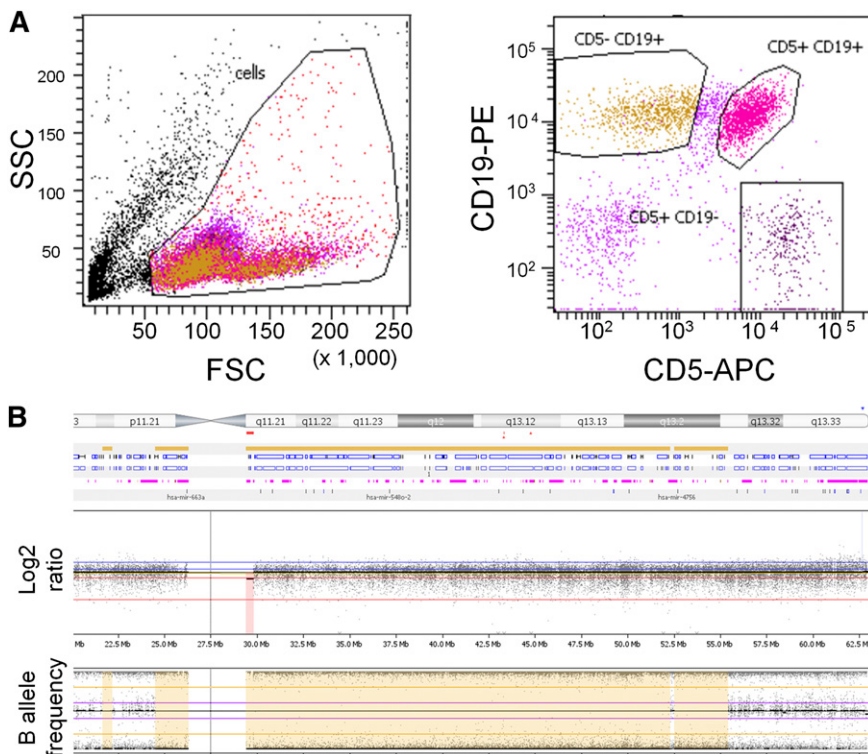


Figure 1. Analysis of an AGS patient. (A) Flow cytometry plot from whole peripheral blood of an AGS patient showing (left) forward-sideward scatter and (right) CD19 PE/CD5 APC scatter gated on lymphoid cells. (B) Whole genome array of chromosome 20 from leukemia cells of an AGS patient, visualized in Nexus (BioDiscovery). The log₂ ratio indicates copy number changes, and the B-allele frequency plot indicates loss of heterozygosity.

analysis was performed. Survival times for the populations of WT *SAMHD1* and mutated *SAMHD1* patients were compared by creating the corresponding Kaplan-Meier survival curves. The statistical difference between the 2 survival curves was quantified by performing the standard log-rank test.³⁵

Results

Initially, we reviewed 32 patients with congenital AGS and germ-line *SAMHD1* mutations. One of these patients was a boy born to consanguineous parents who presented with features typical of AGS.⁷ This patient has a homozygous germ-line canonical splice-acceptor site (1609-1G>C) mutation in intron 14, leading to an alternative splice variant.⁷ Interestingly, at the age of 24 years, he presented with cervical and axillary lymphadenopathy and a lymphocytosis. Flow cytometry demonstrated the typical hallmark of CD19⁺CD5⁺ mature lymphocytes (Figure 1A), confirming CLL. High-resolution genome-wide array analysis of DNA from his CLL cells revealed a large region of homozygosity including the *SAMHD1* locus on chromosome 20 (Figure 1B). This was not associated with a copy number change but was consistent with a germ-line inheritance pattern arising from consanguineous parents carrying shared haplotypes across this region. No other chromosomal aberrations associated with CLL and no novel aberrations were detected.

To establish whether acquired mutations were present, we tested DNA from the patient's CLL cells using a targeted approach to sequence genes reported as mutated recurrently in CLL (supplemental Table 2). None of these genes were mutated in the patient's sample. Next, we performed WES to identify other acquired mutations that might contribute to CLL pathogenesis, including recently described pathogenic mutations⁵ (Materials and methods). Again, no acquired mutations seen previously in CLL or other hematological

malignancies were observed. Although these results do not rule out a contribution of other currently unknown genes in CLL pathogenesis, they support an involvement of *SAMHD1*.

We therefore set out to establish whether acquired *SAMHD1* mutations were recurrent in CLL patients. In our single institution cohort of 100 consecutive CLL patients,³⁶ we identified 5 patients, carrying between them 6 mutations in *SAMHD1* (Figure 2A; Table 1, cohort 1). Germ-line analysis using sorted CD5⁺CD19⁻ cells could be performed for 5 of the mutations and confirmed that they were acquired (supplemental Table 5). Furthermore, 4 of the 5 patients carried a mutation associated with either a mono-allelic deletion or cnLOH as revealed by genome-wide array data (Figure 2A; Table 1) and reflected in the predominantly high VAF (Table 1). Mutations affected mainly conserved residues and were predicted to be deleterious for protein function by in silico prediction (Figure 2B; supplemental Figure 2; supplemental Table 5). One additional patient had a small bi-allelic deletion of the 5' untranslated region of *SAMHD1* (Figure 2C).

To investigate the association between *SAMHD1* mutations and clinical outcome, we analyzed a further 263 UK clinical trials samples by Sanger sequencing and NGS. These patients were from 2 pretreatment trials (cohort 2, n = 200) and 2 relapsed/refractory trials (cohort 3, n = 63; Materials and methods). We found a mutation frequency of 3% in the pretreatment group and of 11% in the relapsed and refractory group (Table 1). Interestingly, the array analyses of 3 patients showed mosaicism, confirming the acquired nature of the copy number change involving the *SAMHD1* locus.

Mutations were distributed across all exons, similar to those identified in AGS (Figure 3A). Three mutations (R145X, R145Q, and I201N) were common to both AGS and CLL patients. When assessed using genome-wide high-resolution arrays, *SAMHD1* mutations showed a clear association with chromosome 20 abnormalities involving the *SAMHD1* locus (seen in 12 of 18 patients; Table 1;

Figure 2. Representative examples of somatic *SAMHD1* mutations. (A) Electropherogram of *SAMHD1* mutations using paired germ-line and tumor DNA. (Upper) Representative results from 2 patients are depicted. (Lower) Whole genome arrays showing (left) copy neutral loss of heterozygosity (cnLOH) or (right) a heterozygous deletion over chromosome 20 involving the *SAMHD1* locus. (B) Multiple sequence alignment of *SAMHD1* around the mutated residue (arrow) of each patient. Degree of conservation is reflected by background color code (dark blue, highly conserved to white, not conserved). (C) Whole genome array showing a small homozygous deletion at the 5' end of the *SAMHD1* locus in patient CLL 095.

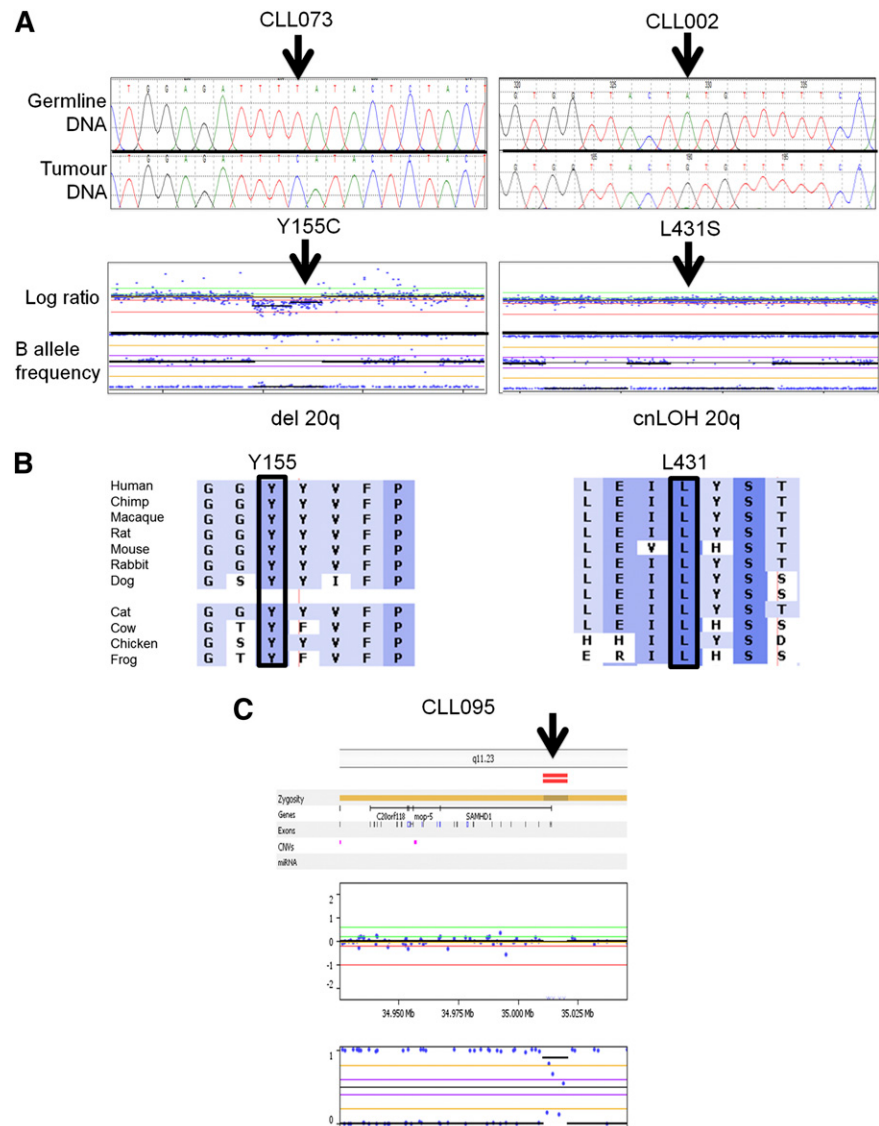


Figure 3B). *SAMHD1* mutations were not found to associate with any other described CLL mutations (Figure 3B; supplemental Table 4). As expected for a group of CLL patients requiring treatment, both treatment naïve and relapsed/refractory patients with *SAMHD1* mutations showed unmutated immunoglobulin (IgVH) rearrangements in the majority of cases. Only 3 of 11 relapsed/refractory patients had del11q compared with 3 of 7 in the pretreatment group (Figure 3B). Together, these data do not support the notion that the enrichment for *SAMHD1* mutations in advanced CLL could be due to a higher prevalence of these confounding poor risk factors in this group.

To evaluate the impact of *SAMHD1* mutations on clinical outcome, we reviewed primary response data on the clinical trial samples (Figure 3B). For the first-line patients recruited into AdMiRe, the end points of CR and bone marrow MRD were analyzed. The CR rate for the AdMiRe cohort, excluding the *SAMHD1*-mutated patients, was 137 of 210 patients (65.2%; 95% exact confidence interval [CI]: 58.4-71.7%).³⁴ The proportion of *SAMHD1*-mutated patients with CR was 0 of 5; this difference in sample proportions was statistically significant (95% CI for difference in proportions: 48.6-81.9%, $P = .014$).

The MRD negative rate across the AdMiRe cohort, excluding the *SAMHD1*-mutated patients, was 108 of 188 patients (57.4%; 95% exact CI: 50.0-64.6%). The proportion of *SAMHD1*-mutated patients who were MRD negative was 0 of 5; this difference in sample proportions was statistically significant (95% CI for difference in proportions: 40.1-74.8%, $P = .035$).

Patients with *SAMHD1* mutations treated in CLL 201 and CLL 202 clinical trials showed outcomes as expected for this cohort of relapsed/refractory patients (Figure 3B). The end point of overall survival was analyzed. The standard log-rank test for comparing the Kaplan-Meier survival curves (supplemental Figure 3) for WT *SAMHD1* and mutated *SAMHD1* samples gave $P = .666$, which suggests that the 2 curves are not significantly different. However, as there were only 6 samples with *SAMHD1* mutations in our overall survival analysis, a larger cohort will be needed to draw more definite conclusions.

SAMHD1 is expressed in normal B cells^{18,37} (Figure 4A). To interrogate the impact of the *SAMHD1* variants, we compared *SAMHD1* mRNA expression in highly purified samples from 11 *SAMHD1* mutated CLL patients with that of normal B cells. Using quantitative RT-PCR, we found statistically significant

Table 1. Summary of *SAMHD1* mutations

Patient number	Cohort	Sample ID	Mutation position	Mutation type	Amino acid	<i>SAMHD1</i> locus (array)	VAF (%)
1	1	CLL077	g.35 526 336 A>T	Missense	F545L	cnLOH	99
2	1	CLL 092	g.35 580 045 T>A	Missense	M1K	cnLOH	72
	1	CLL 092	g.35 526 886 ->A	Frameshift insertion	C522X	cnLOH	27
3	1	CLL073	g.35 563 477 A>G	Missense	Y155C	Loss	94
4	1	CLL 002	g.35 533 885 T>C	Missense	L431S	Loss	92
5	1	CLL 238	g.35 532 562 C>A	Missense	D501Y	Normal	40
6	2	CLL 242	g.35 540 906 G>A	Missense	R371H	Normal	50
7	2	CLL 069	g.35 563 507 G>A	Missense	R145Q	cnLOH	98
8	2	CLL 228	g.35 526 256 G>A	Nonsense	W572X	cnLOH	99
9	2	CLL354	g. 35 532 585 A>C	Missense	L493R	Normal	55
10	2	CLL 007	g. 35 559 171 T>C	Missense	H206R	Mosaic cnLOH	18
	2	CLL 007	g.35 559 186 A>T	Missense	I201N	Mosaic cnLOH	50
11	3	CLL 2537	g.35 540 955 G>A	Missense	E355K	Loss	99
12	3	CLL 2353	g.35 563 508 C>T	Nonsense	R145X	Loss	87
13	3	CLL 2247	g. 35 540 925 T>G	Missense	T365P	Mosaic cnLOH	66
14	3	CLL 2198	g. 35 545 407 T>A	Missense	I300L	Normal	36
	3	CLL 2198	g. 35 563 469 G>A	Missense	P158S	Normal	45
	3	CLL 2198	g. 35 580 001 C>A	Missense	D16Y	Normal	48
15	3	CLL 2219	g. 35 547 889 G>A	Missense	L244F	Normal	39
	3	CLL 2219	g. 35 545 437 G>A	Missense	R290C	Normal	39
16	3	CLL 2250	g. 35 533 826 G>A	Missense	R451C	Normal	57
	3	CLL 2250	g. 35 533 825 C>A	Missense	R451L	Normal	33
17	3	CLL 2462	g. 35 532 564 A>C	Missense	V500G	cnLOH	99

lower expression levels of mRNA in *SAMHD1*-mutated CLL cells (Figure 4A). Where possible, we sequenced cDNA extracted from *SAMHD1*-mutated CLL cells by Sanger sequencing using primers designed to cross intron-exon boundaries. This confirmed that, at least in these patients, *SAMHD1* transcripts in CLL cells consisted exclusively of mutant *SAMHD1* (supplemental Figure 4A). Western blots performed on paired samples also confirmed reduced protein expression compared with WT *SAMHD1* CLL (supplemental Figure 4B).

Furthermore, we assessed *SAMHD1* mRNA (Figure 4B) and protein levels (Figure 4C) in 10 untreated CD19⁺, CD5⁺ samples from consecutive CLL patients compared with healthy control PBMCs and observed a marked reduction of *SAMHD1* mRNA and protein expression in 8 of 10 cases. One case (P9) showed significant mRNA and protein expression, and another case (P10) had no protein expression despite presence of mRNA. Together, these observations suggest that *SAMHD1* mRNA and protein expression in CLL is heterogeneous and that in addition to mutations within the *SAMHD1* coding region itself, other control mechanisms are involved in regulating its expression in CLL.

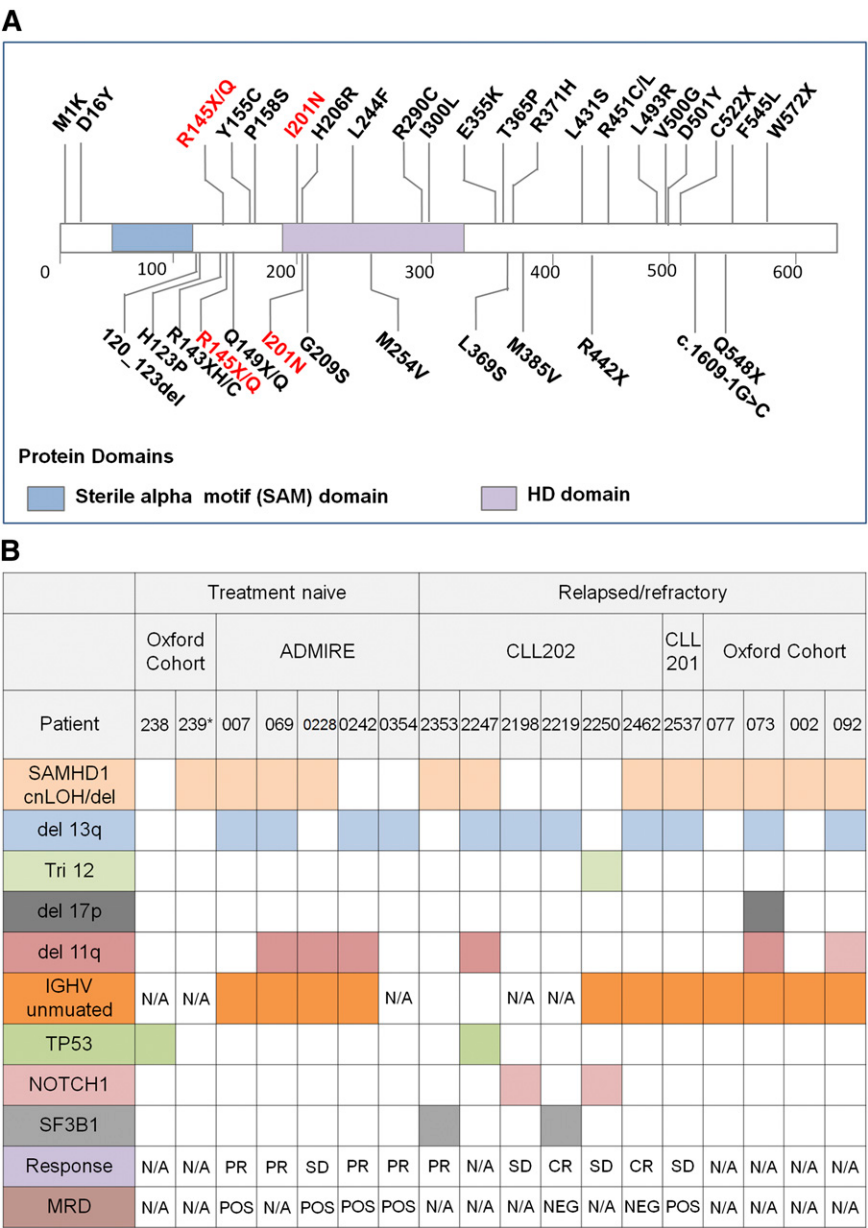
Next, we extended our analysis to other malignancies. *SAMHD1* mutations have been identified in colorectal,³⁸ lung, and urinary tract cancer.³⁸ However, WGS analysis of 1006 different tumor cell lines showed that *SAMHD1* mutations in solid tumors are rare⁴⁰ (supplemental Table 6), and of only 16 cell lines with mutations, 5 were of hemopoietic/lymphoid and 11 of solid tumor origin. *SAMHD1* is expressed in all normal human tissues¹⁷ (supplemental Figure 5). Assessment of *SAMHD1* protein (supplemental Figure 6A) and mRNA (supplemental Figure 6B) expression in different tumor cell lines revealed low level expression in all cell lines tested with the exception of the human acute monocytic leukemia cell line, THP1, known to express normal levels of WT *SAMHD1*.⁹ Interestingly, alternative *SAMHD1* transcripts (supplemental Figure 6C) were identified in all the cell lines tested except in THP1 cells. Additionally, primary breast cancer specimens from 34 individuals demonstrated reduced to absent protein expression of *SAMHD1* in 50% of cases (supplemental Figure 6D).

There is evidence that an imbalance in the dNTP pool produces a hypermutator phenotype and that decreased levels of dNTPs

interfere with proper DNA replication and repair.^{22,23} Therefore, we investigated the potential role of *SAMHD1* in cellular proliferation and DNA damage response. HeLa cells were transduced with lentiviral vector expressing, under the control of an inducible promoter, either WT *SAMHD1*, a dNTP triphosphohydrolase-inactive HD mutant,^{9,13} or a C-terminal deletion mutant that has been characterized as a constitutively active form³³ (supplemental Figure 7). Expression of WT *SAMHD1*, but not HD mutant, resulted in slow proliferation compared with cells transduced with empty vector (Figure 5A). Interestingly, the antiproliferative activity of *SAMHD1* was amplified when the constitutively active mutant (*SAMHD1*ΔCter) was expressed, suggesting a key role for *SAMHD1* enzymatic activity. Next, HeLa cells expressing WT *SAMHD1* were treated with 3 DNA damaging agents, camptothecin, mitomycin C, and etoposide (known to induce DNA double-strand breaks [DSBs]), and with a methylating agent, methyl methanesulfonate. The proportion of living cells was measured after 3 days and compared with baseline (Figure 5B). In this system, WT *SAMHD1* increased cell death after treatment with DNA DSB agents. However, no effect on cell survival was seen with the methylating agent, suggesting that *SAMHD1* is involved specifically in DNA DSB repair. To explore whether *SAMHD1* acquires new function following DNA damage by engaging in specific protein interactions, we analyzed its distribution in nuclear extracts prepared from mock- or etoposide-treated HeLa *SAMHD1* cells using a glycerol gradient sedimentation assay that allows size-dependent separation of protein complexes. We found that *SAMHD1* resides in a high-molecular-weight complex (HMW) in untreated cells and converts to a low-molecular-weight complex after 4-hour treatment with etoposide (Figure 5C). Conversion from the HMW to the low-molecular weight complex is transient, because *SAMHD1* was recovered in the HMW complex 18 hours after etoposide treatment.

We and others have recently uncovered the phosphorylation of *SAMHD1* at threonine 592 (Thr592) by CyclinA2/CDK1 as a key regulatory mechanism of its antiviral activity.^{33,41} Interestingly, we found that treatment of cells with etoposide increased the pool of *SAMHD1* interacting with CyclinA (Figure 5D) and induced *SAMHD1* phosphorylation at the Thr592 (Figure 5E). To further explore the

Figure 3. Schematic representation of *SAMHD1* mutations and relationship to other genetic lesions. (A) Schematic representation of *SAMHD1* protein. (Upper) Mutations found in current study. (Lower) Known mutations in AGS patients. Mutations common to both diseases are shown in red. (B) Associations between *SAMHD1* mutations and other genetic aberrations and clinical outcome. cnLOH, copy neutral loss of heterozygosity; CLL239*, AGS patient; Response, clinical response according to the iwCLL guidelines⁴⁶; PR, partial response; CR, complete response; SD, stable disease; MRD, minimal residual disease; N/A, not available.



involvement of *SAMHD1* in DNA damage response, we determined whether *SAMHD1* is recruited to the site of DNA damage. HeLa cells expressing *SAMHD1* were mock-treated or treated with camptothecin and stained for both 53BP1 and *SAMHD1*. 53BP1 foci were seen in untreated cells presumably reflecting DSBs that resulted from replication stress. An increase in number and size of 53BP1 foci was observed on camptothecin treatment. Interestingly, *SAMHD1* colocalized with 53BP1 foci in both untreated and in CPT-treated cells (Figure 5F). These results suggest that *SAMHD1* is recruited to the site of DNA damage. Taken together, our results show that *SAMHD1* is involved in the response to DNA DSBs and engages in specific protein interactions on DNA damage.

Discussion

In this article, we present a detailed molecular description of patients with *SAMHD1*-mutated CLL recruited into clinical trials

and begin to explore the role of *SAMHD1* in CLL pathogenesis. We focused initially on the molecular characterization of an AGS patient in whom a homozygous germ-line mutation in *SAMHD1* had been identified previously.⁷ Remarkably, this patient developed CLL at only 24 years of age; only 0.2% of patients diagnosed with CLL are 20 to 34 years of age, and the median age at presentation is 72 years.⁴² To our knowledge, there have been no other reports of CLL in AGS patients, although only a small number of patients beyond the age of 20 years are known to us. When we used targeted NGS, WES, and a genome-wide SNP array, approaches to test CLL-derived B cells from this patient, we confirmed the *SAMHD1* mutation status and identified a large region of homozygosity consistent with inheritance from consanguineous parents carrying shared haplotypes. Importantly, we did not identify any acquired CLL-related sequence changes^{2,4,5,30,31} or structural variants^{1,43} in the leukemia cells. These data support the notion that a homozygous *SAMHD1* germ-line mutation was sufficient to cause early-onset CLL in this particular patient. Alternatively, other currently

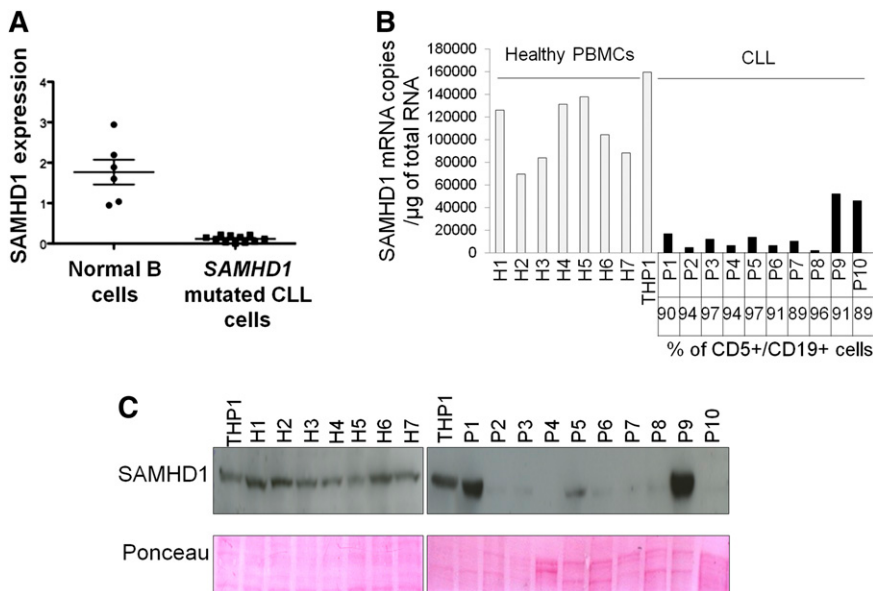


Figure 4. Expression profiles of SAMHD1 in primary CLL cells. (A) cDNA pools from B cells (CD19⁺, CD5⁻) sorted from peripheral blood of healthy donors and SAMHD1-mutated CLL cells (CD19⁺, CD5⁺) were subjected to quantitative PCR analysis using ABL as a control gene. (B) Measurement of SAMHD1 mRNA by quantitative RT-PCR. The percentage of leukemia cells (CD5⁺/CD19⁺) is indicated. Statistical comparison of SAMHD1 mRNA levels between each CLL sample and the mean of the healthy PBMCs was performed. The black bars indicate $P < .01$. (C) Measurement of SAMHD1 protein levels by western blot in 10 CLL samples and 7 healthy PBMCs. Ponceau dye staining was used to control protein load.

unknown acquired or germ-line driver mutations could be involved in the CLL pathogenesis.

The AGS CLL patient findings together with previous longitudinal CLL WGS results in which an acquired homozygous *SAMHD1* mutation was identified led us to perform detailed molecular studies of *SAMHD1* using well-characterized samples from patients recruited into clinical trials. Significantly, our sequencing data revealed that *SAMHD1* mutations occur at a fourfold higher frequency in the relapse/refractory cohort, and the majority are likely deleterious as shown by in silico analyses. Furthermore, the genome-wide SNP array data showed that the mutations in this group associate with cnLOH or monoallelic losses encompassing *SAMHD1*. WGS and WES analyses have been used previously to document recurrent mutations in CLL²⁻⁵ and most recently, *SAMHD1* mutations have been noted at a statistically significant frequency of 4%.⁵ In contrast, extensive WES studies of solid tumor cell lines indicate that genomic mutations of *SAMHD1* in solid tumors are rare events: only 16 mutations were identified in 1006 samples, and this did not reach statistical significance (E. Papaemmanuil, personal communication, January 2013).

Recently, Edelmann et al presented the results of an array-based study of untreated CLL patients recruited into the German CLL8 multicenter trial for first-line treatment.⁴³ In our study, only cohort 2 represented an equivalent pretreatment cohort. For this patient group, the copy number aberration (CNA) findings in our study were consistent with those of the CLL8 study with no CNAs involving *SAMHD1* being detected. For cnLOH, no regions involving *SAMHD1* either in the 144 paired or 209 unpaired analyzed cases were reported in the CLL8 study, whereas we report 3 cases (1.5%) in our cohort 2. All 3 of our cnLOH cases are >10 Mb in size (~14, ~29, and ~33 Mb), but would have been excluded by Edelmann et al⁴³ because they did not coincide with a proven tumor-specific cnLOH or CNA. Additional reasons that may explain why cnLOH regions involving *SAMHD1* were not reported in the CLL8 samples include sample size (direct comparison was only possible for 144 paired CLL8 cases) and differences in the array platform and analysis thresholds used for making cnLOH calls. For example, it is possible that mosaicism, a feature observed in 1 of the 3 cases in our cohort 2, may not have been called in the CLL8 study due to a reduced number of nonhomozygous

intervening SNP calls. Similarly, background noise may have an effect.

To follow-up on our genomic studies, we examined *SAMHD1* mRNA and protein expression in CLL and solid tumors. *SAMHD1* is widely expressed across most normal tissues.^{18,37} Interestingly, although *SAMHD1* mutations were expressed and clearly associated with markedly reduced mRNA and protein levels, reduced expression of *SAMHD1* was also seen in most, but not all, WT *SAMHD1* CLL. One case showed significant mRNA and protein expression, and another case had no protein expression despite presence of mRNA. Similarly, *SAMHD1* mutations were only rarely found in solid tumors, but reduced expression was seen across many cancer cell lines and primary breast cancer. Together, these data suggest (1) that *SAMHD1* mutations might lead to increased degradation of mutant mRNA and/or protein and (2) that other mechanisms apart from mutations must also be responsible for *SAMHD1* regulation. Interestingly, methylation of the *SAMHD1* promoter at CpG island resulting in *SAMHD1* transcriptional repression has recently been reported in CD4⁺ T cell lines such as Jurkat and SupT1.⁴⁴ This further strengthens the idea that a low level of *SAMHD1* expression might involve different regulatory mechanisms.

As a prelude to more detailed functional studies, we presented a series of experiments that provide the first evidence for the potential involvement of *SAMHD1* in the regulation of cell proliferation and response to DNA damage-inducing agents. The regulation of cell proliferation by *SAMHD1* requires its HD domain (Figure 5A). However, our experiments do not determine whether the THP activity and/or the more recently reported nuclease activity of *SAMHD1*¹⁴ are involved. Indeed, the *SAMHD1* HD domain is required for both activities.¹⁴

We showed that *SAMHD1* engages in specific interactions on DNA damage, and its expression regulates cell survival in response to DNA damage-inducing agents. Importantly, we found that *SAMHD1* colocalizes with 53BP1 in DNA repair foci after induction of DSB. Together, these data suggest that *SAMHD1* plays a role in the cellular response to DNA damage. The precise molecular mechanisms underlying these observations are unknown at present. Further studies, including its effect on cellular dNTP availability and genetic experiments of *SAMHD1* knockout models,

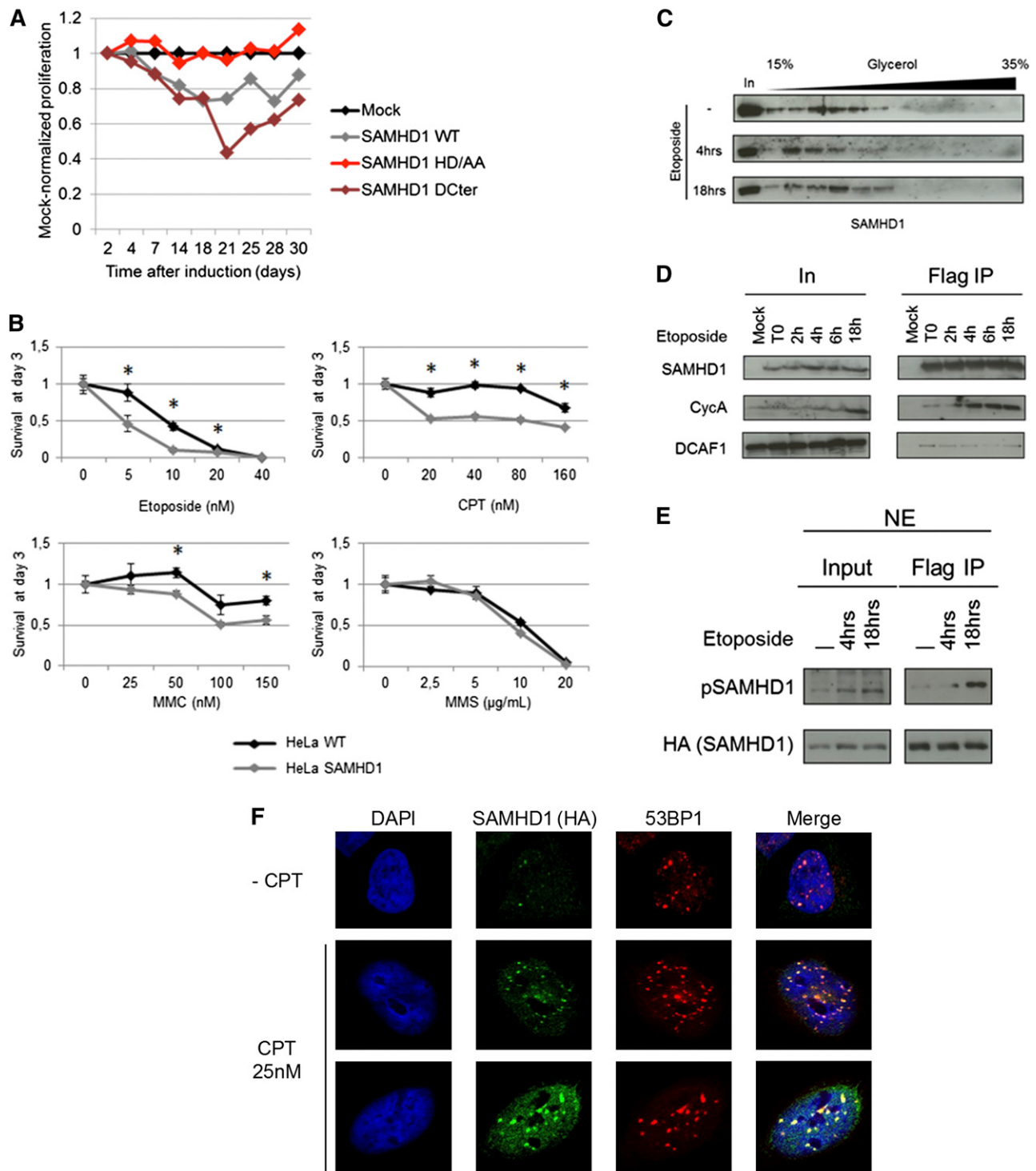


Figure 5. SAMHD1 effects on cell proliferation and cell viability in response to DNA damage agents. (A) HeLa cells stably expressing an inducible empty vector (Mock) or a vector encoding for SAMHD1 wild-type (WT) SAMHD1 mutated in the HD domain (HD/AA) or truncated at amino acid 575 (Δ Cter) were induced with doxycycline. Results are expressed as the proliferation ratio of SAMHD1-transduced cells divided by mock-transduced cell. A representative experiment out of 3 is shown. (B) HeLa and HeLa cells stably expressing SAMHD1 were treated with increasing concentrations of DNA damaging agents, and the percentage of living cells was measured 3 days after treatment. Each graph shows the mean of 3 independent experiments. *A t-test $P < .05$. (C) HeLa cells stably expressing SAMHD1 were mock treated or treated with 5 nM of etoposide for 4 or 18 hours. Nuclear extracts were prepared and separated on a glycerol gradient. The localization of SAMHD1 along the gradient was assessed by western blot. (D) HeLa (Mock) and HeLa cells stably expressing Flag-HA-tagged SAMHD1 were treated with 5 nM etoposide for the indicated time. Flag IPs were performed and SAMHD1 interactions with CycA and DCAF1 were assessed by western blot. (E) Flag IPs were performed using nuclear extract of HeLa cells stably expressing Flag-HA-tagged SAMHD1 after treatment with 5 nM etoposide. SAMHD1 and SAMHD1 phosphorylated on threonine 592 (pSAMHD1) levels were determined by western blot. (F) SAMHD1 colocalizes with 53BP1 at the site of DSBs. HeLa cells expressing SAMHD1-HA were mock-treated or treated with 25 nM CPT. SAMHD1 and 53BP1 intranuclear localizations were determined by immunofluorescence using specific antibodies and confocal microscopy.

will be critical to elucidate further the role of SAMHD1 in DNA damage response.

From a clinical perspective, it is possible that *SAMHD1* mutations in CLL lead to chemotherapy resistance. This is supported by several lines of evidence. (1) Loss of SAMHD1 expression results in enhanced resistance to DSB-inducing agent (Figure 5B). (2) *SAMHD1* mutations are enriched in the relapsed/refractory cohort without apparent increase of poor risk confounding factors such as del11q or unmutated IgVH. (3) The small number of patients with *SAMHD1* mutations in our cohort had a poorer response to first-line chemo-immunotherapy. Alternatively, *SAMHD1* mutations might be associated with disease aggressiveness, defined by a short time to first treatment or transformation to Richter's syndrome. However, in our small number of *SAMHD1*-mutated relapsed/refractory patients, there was no difference in overall survival compared with WT *SAMHD1* patients, suggesting that mutations may not be linked to CLL aggressiveness. Clearly, analyses of larger patient cohorts are required to address this question fully, including also patients with monoclonal B-cell lymphocytosis or Richter's syndrome.

There are important implications of these findings. First, our data illustrate that low-frequency mutations can be candidate drivers of disease and that large collaborative efforts are required to gain insight into the biological and clinical consequences of these low-frequency mutations. Second, mutations in *SAMHD1* are specific to the hematopoietic system and can be founder mutations that occur as early as in the germ-line DNA. This supports their role in the early stages of leukemogenesis. Third, the evidence we present indicates that *SAMHD1* mutations lead to exclusive and reduced expression of mutant mRNA. This poses considerable challenges to the development of drugs targeting the defect. Provided protein function is preserved, 1 possible strategy might be to increase SAMHD1 expression by preventing SAMHD1 degradation.⁴⁵

If our results are confirmed by larger studies, the main implication will be to avoid the use of DNA damaging agents in patients with mutations in *SAMHD1*. In this regard, understanding how SAMHD1 determines the fate of DNA-damaged cell is an important challenge.

In conclusion, we presented recurrent mutations in *SAMHD1* as newly identified molecular events in CLL and demonstrated that the protein SAMHD1 likely engages in specific interactions in response to DNA damage. More work is needed to fully elucidate its role both in regulation of DNA damage response and in leukemogenesis and/or disease progression. Only the analysis of large cohorts of well-characterized clinical trial samples will reveal whether *SAMHD1* mutations emerge as clinically significant response predictors of chemotherapy resistance.

Acknowledgments

The authors thank members of the Molecular Virology laboratory for critical reading of the manuscript and Sabine Laurent-Chabalier for technical assistance.

This work was supported by grants from the European Research Council (250333), Sidaction (Fonds de dotation PIERRE BERGE), Agence Nationale de Recherche sur le Sida, and European FP7 "HIT HIDDEN HIV" contract 305762 (to M.B.). T.L. was supported by Ministère de l'Enseignement supérieur et de la Recherche fellowships. This work was supported by the Oxford Partnership Comprehensive Biomedical Research Centre with funding from the Department of Health's National Institute for Health Research Biomedical Research Centres funding scheme. A.S., A.T.T., J.C.T., and S.J.L.K. are supported also by the Health Innovation Challenge Fund (HICF-1009-026), a parallel funding partnership between the Wellcome Trust and the Department of Health. S.J.L.K. is also supported by the Wellcome Trust Core Award Grant (090532/Z/09/Z). Y.J.C. acknowledges the Manchester Biomedical Research Centre and the Greater Manchester Comprehensive Local Research Network. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant 241779.

The views expressed in this publication are those of the authors and not necessarily those of the Department of Health or the Wellcome Trust.

Authorship

Contribution: R.C., T.L., M.B., and A.S. conceived and designed the experiments and prepared the manuscript; R.C., T.L., A.B., P.R., A.T.T., H.D., J.M., J.B., M.R., Y.-L.L., P.P., and A.T. performed experiments; S.A., P.H., F.S., A.P., and J.G.J. provided samples; R.C., T.L., M.B., A.S., G.W.C., M.T., D.R.C., S.J.H., M.T.R., S.J.L.K., J.C.T., J.R., D.B., and Y.J.C., analyzed data; and A.S. and M.B. oversaw all of the work.

Conflict-of-interest disclosure: M.T.R. and D.B. are employees of Illumina Inc., a public company that develops and markets systems for genetic analysis. The remaining authors declare no competing financial interests.

Correspondence: Anna Schuh, Oxford NIHR Biomedical Research Centre, Molecular Diagnostic Laboratory, Level 4 John Radcliffe Site, University of Oxford, Oxford, UK; e-mail: anna.schuh@ndcls.ox.ac.uk; or Moncef Benkirane, Institut de Genetique Humaine, CNRS UPR1142, Laboratoire de Virologie Moléculaire, Montpellier, France; e-mail: moncef.benkirane@igh.cnrs.fr.

References

- Döhner H, Stilgenbauer S, Benner A, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med*. 2000; 343(26):1910-1916.
- Quesada V, Conde L, Villamor N, et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet*. 2012;44(1):47-52.
- Wang L, Lawrence MS, Wan Y, et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N Engl J Med*. 2011;365(26):2497-2506.
- Puente XS, Pinyol M, Quesada V, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2011; 475(7354):101-105.
- Landau DA, Carter SL, Stojanov P, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*. 2013;152(4):714-726.
- Schuh A, Becq J, Humphray S, et al. Monitoring chronic lymphocytic leukemia progression by whole genome sequencing reveals heterogeneous clonal evolution patterns. *Blood*. 2012;120(20):4191-4196.
- Rice GI, Bond J, Asipu A, et al. Mutations involved in Aicardi-Goutières syndrome implicate SAMHD1 as regulator of the innate immune response. *Nat Genet*. 2009;41(7):829-832.
- Crow YJ, Rehwinkel J. Aicardi-Goutières syndrome and related phenotypes: linking nucleic acid metabolism with autoimmunity. *Hum Mol Genet*. 2009;18(R2):R130-R136.
- Laguerre N, Sobhian B, Casartelli N, et al. SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature*. 2011;474(7353):654-657.

10. Baldauf HM, Pan X, Erikson E, et al. SAMHD1 restricts HIV-1 infection in resting CD4(+) T cells. *Nat Med*. 2012;18(11):1682-1687.
11. Hrecka K, Hao C, Gierszewska M, et al. Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature*. 2011;474(7353):658-661.
12. Descours B, Cribier A, Chable-Bessia C, et al. SAMHD1 restricts HIV-1 reverse transcription in quiescent CD4(+) T-cells. *Retrovirology*. 2012; 9:87.
13. Goldstone DC, Ennis-Adeniran V, Hedden JJ, et al. HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. *Nature*. 2011;480(7377):379-382.
14. Beloglazova N, Flick R, Tchigvintsev A, Brown G, Popovic A, Nocek B, Yakunin AF. Nuclease activity of the human SAMHD1 protein implicated in the Aicardi-Goutieres syndrome and HIV-1 restriction. *J Biol Chem*. 2013;288(12):8101-8110.
15. Powell RD, Holland PJ, Hollis T, Perrino FW. Aicardi-Goutieres syndrome gene and HIV-1 restriction factor SAMHD1 is a dGTP-regulated deoxynucleotide triphosphohydrolase. *J Biol Chem*. 2011;286(51):43596-43600.
16. Lahouassa H, Daddacha W, Hofmann H, et al. SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nat Immunol*. 2012;13(3):223-228.
17. Li N, Zhang W, Cao X. Identification of human homologue of mouse IFN-gamma induced protein from human dendritic cells. *Immunol Lett*. 2000; 74(3):221-224.
18. Descours B, Cribier A, Chable-Bessia C, et al. SAMHD1 restricts HIV-1 reverse transcription in quiescent CD4(+) T-cells. *Retrovirology*. 2012;9:87.
19. Laguette N, Benkirane M. How Samhd1 changes our view of viral restriction. *Trends Immunol*. 2011; (Dec):14.
20. Hollenbaugh JA, Gee P, Baker J, et al. Host factor SAMHD1 restricts DNA viruses in non-dividing myeloid cells. *PLoS Pathog*. 2013;9(6):e1003481.
21. Gramberg T, Kahle T, Bloch N, et al. Restriction of diverse retroviruses by SAMHD1. *Retrovirology*. 2013;10:26.
22. Bester AC, Roniger M, Oren YS, et al. Nucleotide deficiency promotes genomic instability in early stages of cancer development. *Cell*. 2011;145(3): 435-446.
23. Chabosseau P, Buhagiar-Labarchède G, Onclercq-Delic R, Lambert S, Debatisse M, Brison O, Amor-Guérét M. Pyrimidine pool imbalance induced by BLM helicase deficiency contributes to genetic instability in Bloom syndrome. *Nat Commun*. 2011;2:368.
24. Illumina Inc. BaseSpace. <https://basespace.illumina.com>. Accessed June 14, 2012.
25. Lunter G, Goodson M. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res*. 2011; 21(6):936-939.
26. Rimmer A, Phan H, Matieson I, et al. Platypus. <http://www.well.ox.ac.uk/platypus>. Accessed August 2012.
27. Castor A, Nilsson L, Astrand-Grundström I, et al. Distinct patterns of hematopoietic stem cell involvement in acute lymphoblastic leukemia. *Nat Med*. 2005;11(6):630-637.
28. National Centre for Biotechnology Information. dbSNP. <http://www.ncbi.nlm.nih.gov/SNP/>. Accessed August 2012.
29. Abecasis GR, Auton A, Brooks LD, et al; 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491(7422):56-65.
30. Skowronska A, Parker A, Ahmed G, et al. Biallelic ATM inactivation significantly reduces survival in patients treated on the United Kingdom Leukemia Research Fund Chronic Lymphocytic Leukemia 4 trial. *J Clin Oncol*. 2012;30(36):4524-4532.
31. Zenz T, Kröber A, Scherer K, et al. Monoallelic TP53 inactivation is associated with poor prognosis in chronic lymphocytic leukemia: results from a detailed genetic characterization with long-term follow-up. *Blood*. 2008;112(8):3322-3329.
32. Laguette N, Rahm N, Sobhian B, et al. Evolutionary and functional analyses of the interaction between the myeloid restriction factor SAMHD1 and the lentiviral Vpx protein. *Cell Host Microbe*. 2012;11(2):205-217.
33. Cribier A, Descours B, Valadão AL, Laguette N, Benkirane M. Phosphorylation of SAMHD1 by cyclin A2/CDK1 regulates its restriction activity toward HIV-1. *Cell Rep*. 2013;3(4):1036-1043.
34. Hillmen P. Results of the randomised Phase II NCRI ADMIRE trial of FCR and FCM-R in previously untreated CLL: oral FCR is highly effective and safe but the addition of mitoxantrone does not improve responses. *Haematologica*. 2013;98(s1):473.
35. Bewick V, Cheek L, Ball J. Statistics review 12: survival analysis. *Crit Care*. 2004;8(5):389-394.
36. Knight SJ, Yau C, Clifford R, et al. Quantification of subclonal distributions of recurrent genomic aberrations in paired pre-treatment and relapse samples from patients with B-cell chronic lymphocytic leukemia. *Leukemia*. 2012;26(7): 1564-1575.
37. Coon S, Wang D, Wu L. Polymorphisms of the SAMHD1 gene are not associated with the infection and natural control of HIV type 1 in Europeans and African-Americans. *AIDS Res Hum Retroviruses*. 2012;28(12):1565-1573.
38. Seshagiri S, Stawiski EW, Durinck S, et al. Recurrent R-spondin fusions in colon cancer. *Nature*. 2012;488(7413):660-664.
39. Forbes SA, Bindal N, Bamford S, et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res*. 2011;39(Database issue): D945-D950.
40. Makishima H, Yoshida K, Nguyen N, et al. Somatic SETBP1 mutations in myeloid malignancies. *Nat Genet*. 2013;45(8):942-946.
41. White TE, Brandariz-Núñez A, Valle-Casuso JC, et al. The retroviral restriction ability of SAMHD1, but not its deoxynucleotide triphosphohydrolase activity, is regulated by phosphorylation. *Cell Host Microbe*. 2013;13(4):441-451.
42. National Cancer Institute. Seer Cancer Statistics, NCI. <http://archive.is/TKAq>. Accessed November 2013.
43. Edelmann J, Holzmann K, Miller F, et al. High-resolution genomic profiling of chronic lymphocytic leukemia reveals new recurrent genomic alterations. *Blood*. 2012;120(24): 4783-4794.
44. de Silva S, Hoy H, Hake TS, Wong HK, Porcu P, Wu L. Promoter methylation regulates SAMHD1 gene expression in human CD4+ T cells. *J Biol Chem*. 2013;288(13):9284-9292.
45. Wei W, Guo H, Liu X, et al. A first-in-class NAE inhibitor, MLN4924, blocks lentiviral infection in myeloid cells by disrupting neddylation-dependent Vpx-mediated SAMHD1 degradation [published online ahead of print October 23, 2013]. *J Virol*. doi:10.1128/JVI.02568-13.
46. Hallek M, Cheson BD, Catovsky D, et al; International Workshop on Chronic Lymphocytic Leukemia. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood*. 2008;111(12):5446-5456.

Supplementary Material

Supplementary Methods:

Samples

Trial descriptions: AdMIRe is a randomized Phase II Trial of fludarabine, cyclophosphamide and rituximab (FCR) with or without mitoxantrone (M) in previously untreated physically fit patients with CLL. ARCTIC is a randomized phase IIB trial in previously untreated physically fit patients with CLL to compare fludarabine, cyclophosphamide and rituximab (FCR) with FC, mitoxantrone and low dose rituximab (FCM-miniR).

CLL201 is a randomised phase II multicentre trial to assess the efficacy and safety of fludarabine, cyclophosphamide and mitoxantrone in combination with rituximab (FCM-R), and fludarabine, cyclophosphamide and mitoxantrone alone (FCM) in previously treated but not fludarabine-refractory patients with CLL. CLL202 is a phase II non-randomised study of subcutaneous alemtuzumab plus fludarabine in fludarabine-refractory patients. Refractoriness was defined as no response or relapse within 6 months from fludarabine treatment.

Supplementary Figure Legends

Supplementary Figure 1: Whole Genome Sequencing reveals homozygous *SAMHD1* mutation

Whole Genome Sequencing (WGS) on CLL patient CLL077 over 5 treatment timepoints (a-e). All mutations identified by were clustered according to their variant allele frequency (VAF). Mutations with VAF= $\geq$ 50% are depicted. Red lines: all mutations (including intronic mutations) except for non-synonymous mutations. Black lines: non-synonymous mutations.

Supplementary Figure 2: Conservation tracks for all *SAMHD1* mutations

Multiple Sequence Alignment of *SAMHD1* around the mutated residue (arrow) of each patient. Degree of conservation is reflected by background colour code (dark blue: highly conserved to white: not conserved).

Supplementary Figure 3: Kaplan Meier Estimates of Overall Survival

Overall survival of *SAMHD1* mutated and *SAMHD1* wild type (wt) patients in relapsed/refractory clinical trials. *SAMHD1* mutant n=6, *SAMHD1* wt n= 55; p = 0.666 using the standard log rank test.

Supplementary Figure 4: mRNA and Protein Expression of mutant SAMHD1

a Electropherogram of cDNA obtained from *SAMHD1* mutated CLL cells from 3 patients using primers that cover intron/exon boundaries of *SAMHD1*. Arrows indicate mutation compared to reference sequence.

b Western Blot of *SAMHD1* expression in *SAMHD1* mutated, wt *SAMHD1* CLL and ThP1= monocytic cell line as control. RNA expression data (relative values calibrated against normal B cell control expression of 1) for *SAMHD1* mutated samples (see Methods).

Supplementary Figure 5 *SAMHD1* expression in different normal tissue types including epithelial cells. Gene expression data for *SAMHD1* extracted from CCLE_Expression_Entrez_2012-10-18.res:Gene-centric RMA-normalized mRNA expression data.

Supplementary Figure 6 Expression data of *SAMHD1* in solid tumours

- a, b** Protein extracts and mRNA were prepared from the indicated cell lines and the expression level of *SAMHD1* was assessed by Western Blot (a) and by quantitative RT-PCR (b)
- c** RNAs from different cell lines were also subjected to RT-PCR analysis using *SAMHD1* specific oligonucleotides which amplify the full length open reading frame and the PCR products were analyzed on agarose gel

d Western Blot of SAMHD1 using protein extracts from primary solid mammary tumours. SAMHD1 expressing (THP1) and non-expressing (SupT1) cell lines were used as control. Actin was probed as loading control.

Supplementary Figure 7: SAMHD1 expression in transduced HeLa Cells

HeLa cells were transduced with lentiviral vector expressing, under the control of an inducible promoter, either wild-type (wt SAMHD1, a dNTP TPH-inactive HD mutant (HD/AA) or a C-terminal deletion mutant (Δ Cter).

Supplementary Table 1: Next generation sequencing panel

Gene	Chr	Position (bp; hg19)	Targets	# Amplicons
SF3B1	2	198256648 - 198299821	Exons 14, 15, 16 & 18	8
XPO1	2	61704984 - 61765761	Exon 15	2
MYD88	3	38182608 - 38182792	Exon 5	1
IRF4	6	391689 - 411493	Exons 2 & 3	4
ZFPM2	8	106331097 - 106816817	All Exons	33
NOTCH1	9	139390508 - 139392025	Exon 34	9
ATM	11	108093509 - 108239876	All Exons	113
BIRC3	11	102188144 - 102208515	All Exons	38
TP53	17	7571670 - 7590913	All Exons	21
SAMHD1	20	35520177 - 35580296	All Exons	26
MED12	X	70338356 - 70362354	All Exons	78

Supplementary Table 2: Primers for DNA Sanger Sequencing

Gene_Exon	Forward Primer	Reverse Primer
SAMHD1_exon1	TTGAGGACGACTGGACTGC	GCCTCGGGTCTTCCTTTC
SAMHD1_exon2	GGATCTGGGTAAATGTTGGTG	GGCTTTGTCCCTGAAAGATG
SAMHD1_exon3	TTGTCCTTCCAATCTAACTTCTG	TCACTGAGAAGCAGATTTCCTC
SAMHD1_exon4	AATTTCAAGTTGGACTTACATGTCG	CCATGCCTGGCCTAAGATAAC
SAMHD1_exon5	TTGAGGACCAGAACTATCACTCC	TTCCATATTCTCTTGTTGATCTG
SAMHD1_exon6	TGTTGGTTTACAGATGTTTTCTTG	GCACCCTGGACACTGTAATG
SAMHD1_exon7	GTTAGTCAGGGCTCCCAATG	GGCTAGATGAAAAGCACCAGTC
SAMHD1_exon8/9	TGCCTATCCTAAAACTTCCAGTG	TCCAAATTTAGACCAGGAAC
SAMHD1_exon10	TGCATAGCCCTCTTGCTCTGC	GCCAAGTATCAAGGAAAAATTTATAGG
SAMHD1_exon11	TGTGCCTCATTATTTGTTTCC	ATTGGCAATTCAGGGACTTC
SAMHD1_exon12	TTTGCGAACTGCCTGTTAAG	CCTCTTGGAGGACAGAGATAACC
SAMHD1_exon13	TGTTTGTGGCTCAAAGACTTG	TGGGTGCTTTATCTTTAAACG
SAMHD1_exon14/15	GGCCTCTGTTCCATTCTTG	GTGCAAAGTTTGTGAGTAACAGG
SAMHD1_exon16	TGTGCGCACTTGAGACTTTC	GTGCAAAGTTTGTGAGTAACAGG

Supplementary Table 3: Primers for cDNA Sanger Sequencing

Gene_Exon	Forward Primer	Reverse Primer
SAMHD1_exon12	AGCAATTGACGACATGGAAG	CTGGCAACCTCTTTTGAAG
SAMHD1_exon15	TAAGACTGCCCCCAACAGAG	TGAGTGGGGCTATAACATCG

Supplementary Table 4: Results of association analysis of *SAMHD1* mutated and wild type samples with common CLL-associated mutations

Genetic lesion	p value
del17p13	0.869
TP53	0.475
del11q22	0.048
trisomy12	0.374
del13q14	0.774
NOTCH1	0.809
SF3B1	0.237

Supplementary Table 5: Pathogenicity scores and Germline confirmation of *SAMHD1* mutations

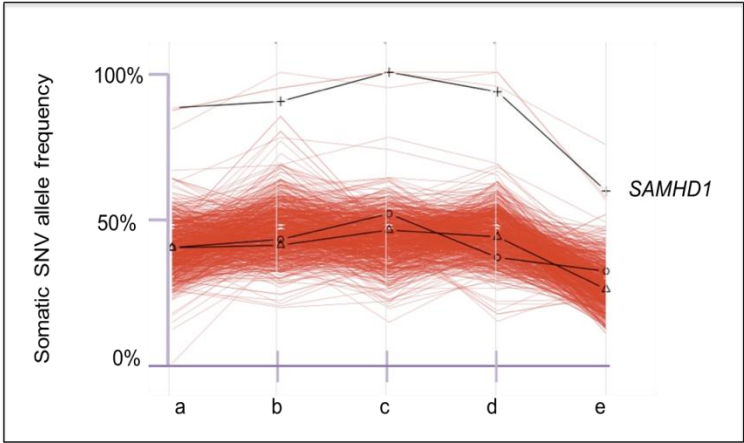
SIFT¹ and Polyphen²

Patient no	Cohort	Sample ID	Amino acid change	SIFT prediction	SIFT score	Polyphen-2 Prediction	Polyphen Score	Germline confirmation
1	1	CLL077	F545L	T	0.54	D	0.978	Y
2	1	CLL 092	M1K	U	-	P	0.915	Y
	1	CLL 092	C522X	U	-	U	-	Y
3	1	CLL073	Y155C	D	0.05	D	0.999	NA
4	1	CLL 002	L431S	D	0.00	D	0.992	Y
5	1	CLL 238	D501Y	T	0.07	P	0.655	Y
6	2	CLL 242	R371H	T	0.09	D	0.999	NA
7	2	CLL 069	R145Q	D	0.02	D	0.999	NA
8	2	CLL 228	W572X	U	-	U	-	NA
9	2	CLL354	L493R	D	0.01	B	0.446	NA
10	2	CLL 007	H206R	D	0.01	D	0.999	NA (mosaic)
	2	CLL 007	I201N	D	0.00	D	1.0	NA (mosaic)
11	3	CLL 2537	E355K	T	0.20	P	0.871	N/A
12	3	CLL 2353	R145X	U	-	U	-	Y
13	3	CLL 2247	T365P	D	0.02	D	0.999	NA (mosaic)
14	3	CLL 2198	I300L	D	0.03	D	0.995	NA
	3	CLL 2198	P158S	D	0.02	D	1.000	NA
	3	CLL 2198	D16Y	D	0.00	D	1.000	NA
15	3	CLL 2219	L244F	D	0.03	D	0.993	NA
	3	CLL 2219	R290C	T	0.12	D	1.000	NA
16	3	CLL 2250	R451C	D	0.01	D	1.000	NA
	3	CLL 2250	R451L	D	0.02	D	0.999	NA
17	3	CLL 2462	V500G	D	0.00	D	0.999	NA

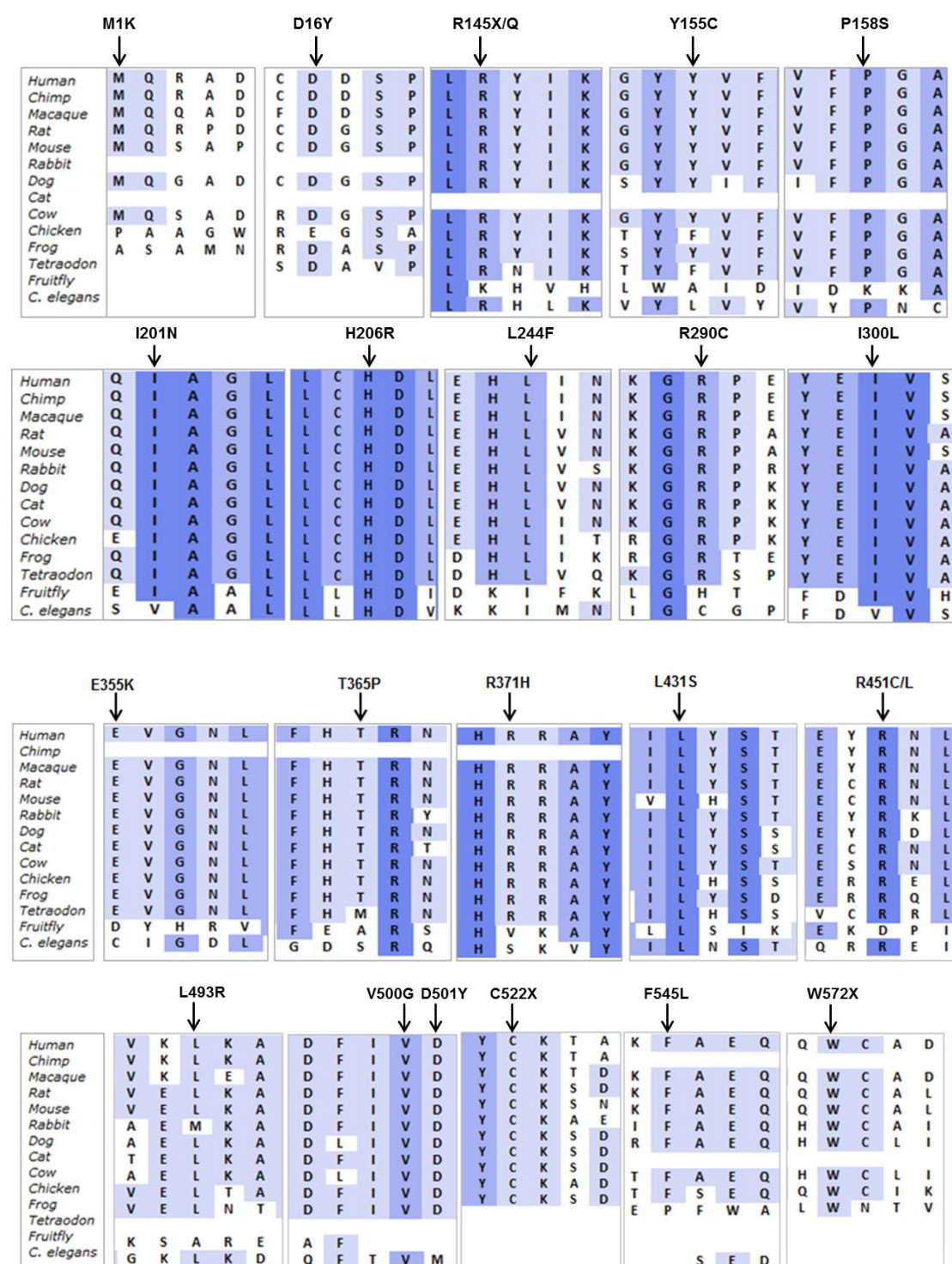
Supplementary Table 6: SAMHD1 mutations in cell lines

Cell Line	Tumour type	Coordinates	Mutation type	Nucleotide position	Amino acid substitution	SIFT	Polyphen-2
J-RT3-T3-5	Lymphoid neoplasm	chr20:35521385-35521385	nonsense	c.1831C>T	p.R611*	untested	untested
TGBC11TKB	Stomach carcinoma	chr20:35521462-35521462	missense	c.1754A>G	p.D585G	deleterious	benign
MFE-319	Endometrial adenocarcinoma	chr20:35526931-35526931	missense	c.1520A>G	p.Y507C	deleterious	probably damaging
LNCaP-Clone-FGC	Prostate carcinoma	chr20:35532579-35532579	missense	c.1484C>T	p.A495V	tolerated	probably damaging
GR-ST	Haematopoietic neoplasm	chr20:35533825-35533825	missense	c.1352G>A	p.R451H	deleterious	probably damaging
REH	Lymphoid neoplasm	chr20:35533825-35533825	missense	c.1352G>A	p.R451H	deleterious	probably damaging
KARPAS-422	Lymphoid neoplasm	chr20:35533834-35533834	missense	c.1343T>C	p.I448T	deleterious	benign
SK-N-DZ	Neuroblastoma	chr20:35533852-35533852	missense	c.1325G>C	p.R442P	deleterious	probably damaging
NCI-H1155	Lung carcinoma	chr20:35533856-35533856	missense	c.1321G>A	p.A441T	deleterious	probably damaging
SJRH30	Rhabdomyosarcoma	chr20:35545171-35545171	missense	c.1016G>A	p.R339H	deleterious	probably damaging
SHP-77	Lung carcinoma	chr20:35545440-35545440	missense	c.865G>A	p.G289R	deleterious	probably damaging
SK-MEL-28	Malignant melanoma	chr20:35555584-35555584	essential splice	c.696+1G>A	p.?	untested	untested
TUR	Lymphoid neoplasm	chr20:35555634-35555635	deletion	c.646_647delAT	p.M216fs*2	untested	untested
PC-3_JPC-3	Lung adenocarcinoma	chr20:35563454-35563454	missense	c.487A>G	p.N163D	deleterious	probably damaging
NCI-H2228	Lung carcinoma	chr20:35575208-35575208	essential splice	c.209-1G>C	p.?	untested	untested
NCI-H1105	Lung carcinoma	chr20:35580028-35580028	nonsense	c.19G>T	p.E7*	untested	untested

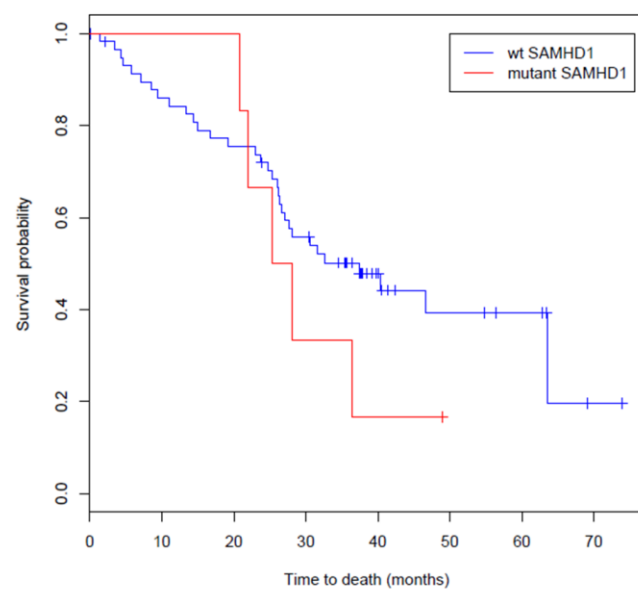
Supplementary Figure 1



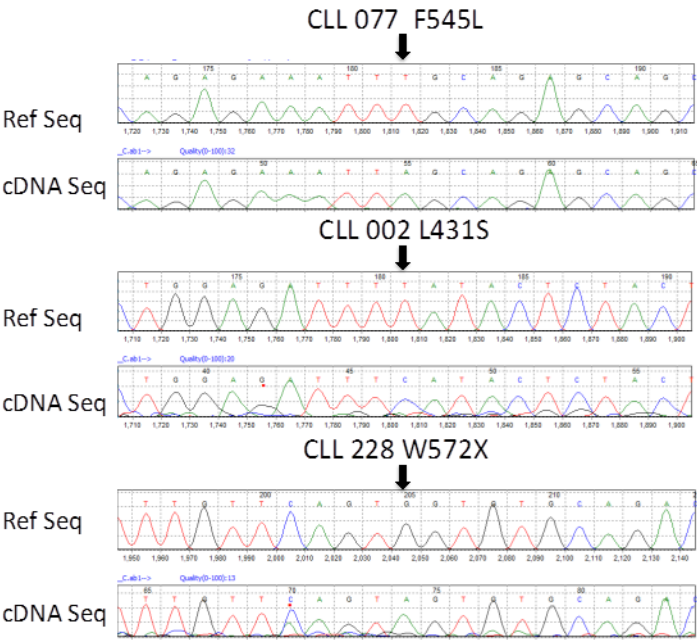
Supplementary Figure 2



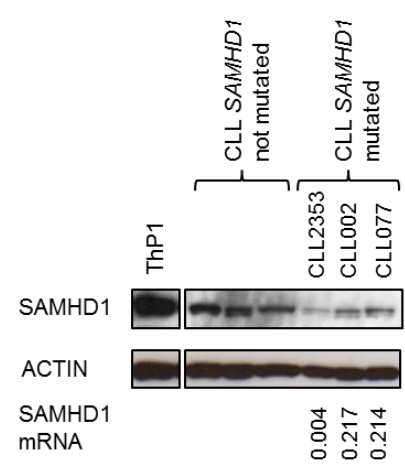
Supplementary Figure 3



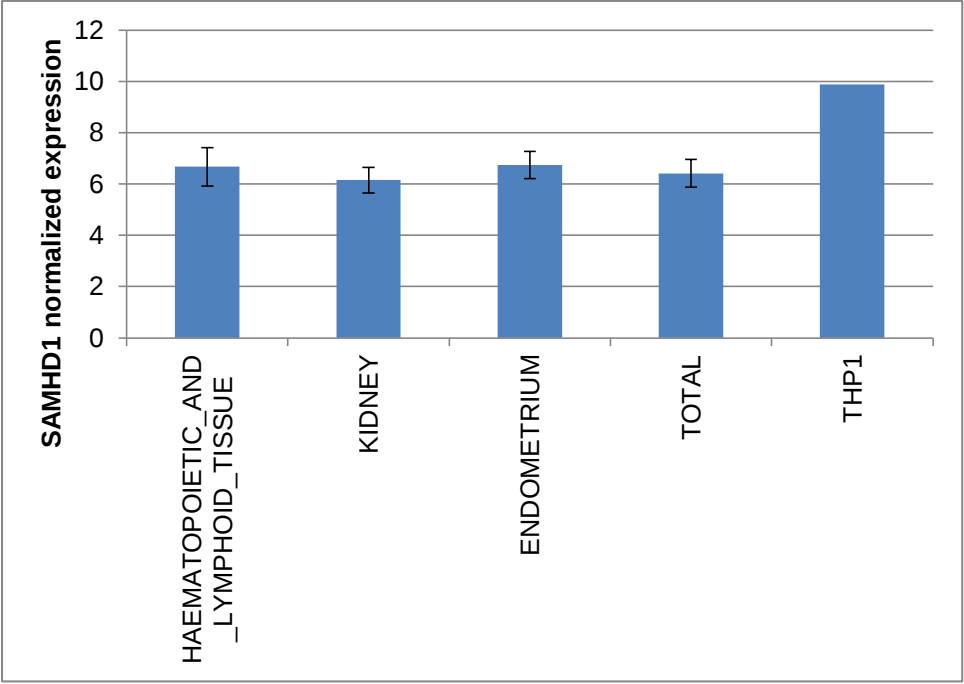
Supplementary Figure 4a



Supplementary Figure 4b

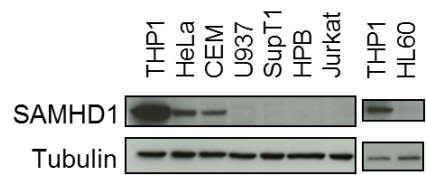


Supplementary Figure 5

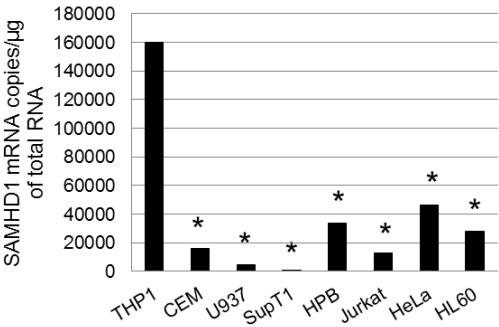


Supplementary Figure 6

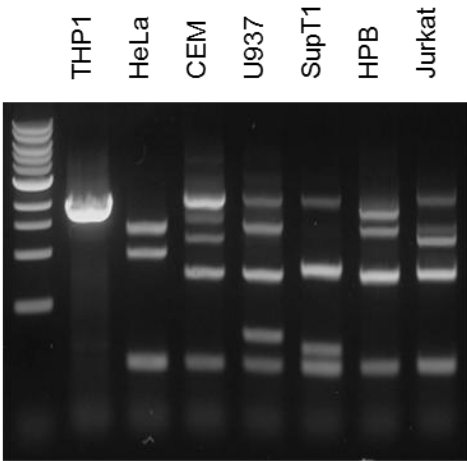
a



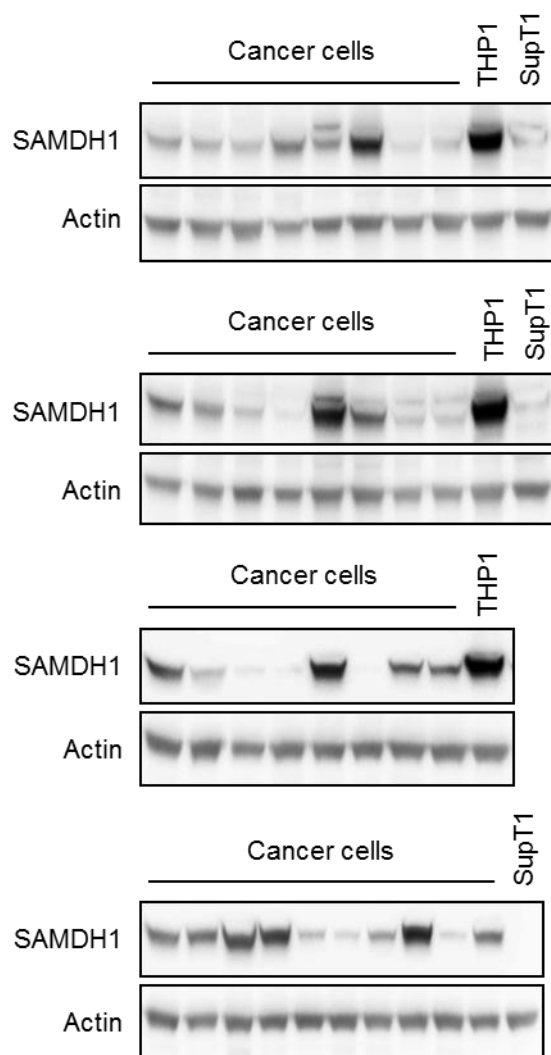
b



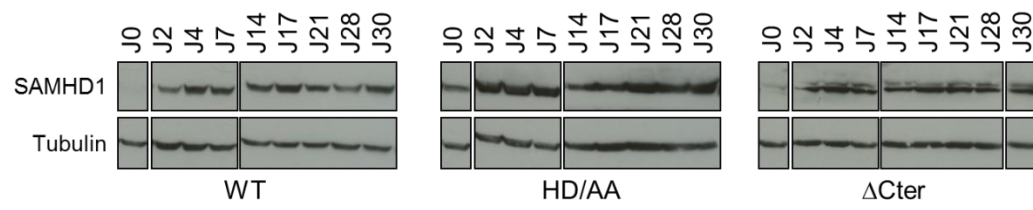
c



d



Supplementary Figure 7



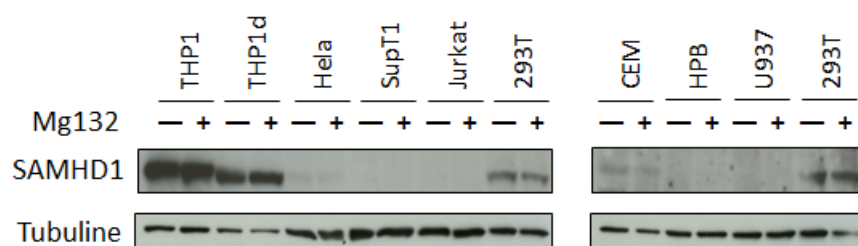
References

1. Kumar, P., Henikoff, S. & Ng, P.C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* **4**, 1073-81 (2009).
2. Adzhubei, I.A. *et al.* A method and server for predicting damaging missense mutations. *Nat Methods* **7**, 248-9 (2010).

Présentation de ma démarche expérimentale

Sachant que SAMHD1 est une dNTP triphosphohydrolase et que les dNTP sont essentiels à différents processus cellulaires comme la réplication du génome dans les cellules en prolifération et la réparation des dommages à l'ADN, j'ai voulu savoir si SAMHD1 pouvait avoir un impact sur le déroulement de ces processus.

Pour cela, j'ai commencé par vérifier le niveau d'expression de SAMHD1 dans différentes lignées utilisées au laboratoire afin de choisir un modèle pertinent pour la suite des expérimentations. Il était déjà connu que la lignée cellulaire THP-1 exprime SAMHD1 à un niveau comparable à ceux observés dans des cellules hématopoïétiques primaires (Laguette et al., 2011), j'ai donc utilisé cette lignée comme contrôle auquel comparer les autres lignées. J'ai constaté que dans l'ensemble des lignées étudiées, l'expression de SAMHD1 est réduite par rapport à celle observée dans les THP-1 (Supplementary Figure 6A). Afin de voir si ce défaut d'expression était dû à une différence dans la stabilité de la protéine ou à une différence dans sa production, j'ai traité les lignées utilisées avec un inhibiteur du protéasome et mesuré l'expression de SAMHD1 par immunoblot (Résultat additionnel 1). L'inhibiteur du protéasome ne stabilisant l'expression de SAMHD1 dans aucune des lignées considérées, j'en ai déduit que la réduction du niveau d'expression de SAMHD1 était dû à une réduction de la production de cette protéine.

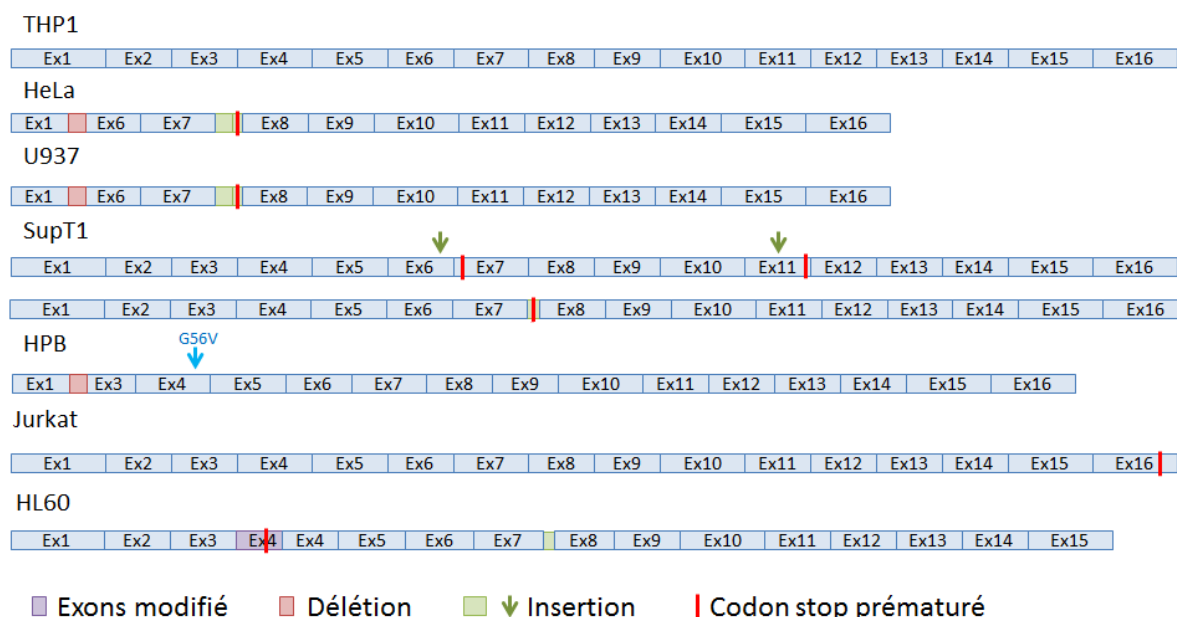


Résultat additionnel 1 : inhiber le protéasome ne permet pas de stabiliser SAMHD1 dans les lignées cellulaires. Les lignées cellulaires indiquées ont été traitées avec du Mg132 10 μ M pendant 4 heures. Les extraits protéiques totaux ont ensuite été purifiés et l'expression de SAMHD1 a été évaluée par immunoblot.

J'ai par la suite quantifié le niveau d'expression des ARNm SAMHD1 par PCR quantitative (qPCR) dans ces différentes lignées (Supplementary Figure 6B) et j'ai constaté que la réduction de l'expression de SAMHD1 par rapport aux cellules THP-1 était déjà présente au niveau des ARNm. Je me suis donc demandé si, en plus de leur quantité, il y avait des différences entre les ARNm SAMHD1 transcrits dans les différentes lignées cellulaires. J'ai donc réalisé des PCR sur des ADNc (complémentaires) purifiés à partir des différentes lignées, en utilisant des amorces localisées dans les régions terminales non-traduites (UTR) du gène SAMHD1. J'ai alors constaté que, si un seul transcrit correspondant à l'ARNm sauvage était amplifié dans la lignée THP-1, plusieurs transcrits différents les uns des autres et de tailles variables pouvaient être observés dans les autres lignées (Supplementary Figure 6C). La plupart de ces transcrits sont plus courts que le transcrit sauvage, mais aucune protéine tronquée n'a été détectée par immunoblot, malgré l'utilisation de trois anticorps différents.

Afin d'identifier les variations de séquence entre les différents transcrits, pour chaque lignée cellulaire le transcrit dont la taille se rapprochait le plus de celle du transcrit sauvage a été sous-cloné et séquencé (Résultat additionnel 2). La séquence des transcrits est mutée dans l'ensemble des lignées considérées à l'exception des THP-1. La plupart de ces mutations sont à l'origine de perturbation dans

l'épissage des transcrits SAMHD1 et certaines d'entre elles aboutissent à la création de codons stop prématurés.



Résultat additionnel 2 : SAMHD1 est muté dans les lignées cellulaires. Des ADNc SAMHD1 de taille proche du sauvage ont été sous-clonés et séquencés. Les séquences obtenues sont ici schématisées.

L'ensemble de ces résultats montre que l'expression du gène SAMHD1 est perturbée dans la plupart des lignées cellulaires utilisées en laboratoire. Cette observation est en accord avec notre hypothèse initiale (depuis également renforcée par d'autres observations) d'une influence de SAMHD1 sur le cycle cellulaire. En effet, les lignées utilisées en laboratoire sont des lignées transformées ayant acquis la capacité de se diviser un plus grand nombre de fois et plus rapidement que les cellules primaires d'origine. SAMHD1 étant une dNTPase réduisant la quantité de dNTP disponibles, cette protéine pourrait constituer un frein au processus de réplication du génome nécessaire avant la division cellulaire et sa mutation pourrait favoriser la prolifération des lignées utilisées en laboratoire.

Afin de vérifier cette hypothèse, j'ai étudié l'effet de SAMHD1 sur le cycle cellulaire. Pour des raisons expérimentales, j'ai choisi de surexprimer SAMHD1 dans une lignée où sa forme sauvage n'est pas transcrite (la lignée HeLa) plutôt que de supprimer son expression dans la lignée THP-1.

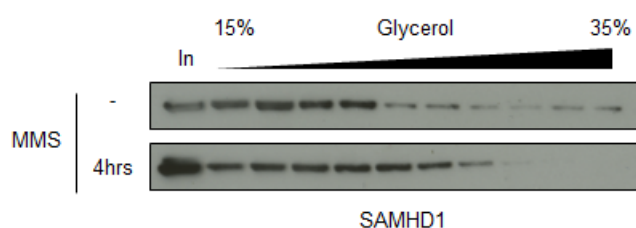
J'ai transduit des cellules HeLa avec des vecteurs permettant l'expression de variants de SAMHD1 sous le contrôle d'un promoteur inductible pour obtenir quatre lignées : une lignée contrôle, une pouvant exprimer SAMHD1 sauvage, une pouvant exprimer le mutant HD/AA dont le site catalytique est inactif et une pouvant exprimer le mutant DeltaCter tronqué au niveau de l'acide aminé 575 et n'exprimant donc plus la T592. La phosphorylation sur cette T592 inhibant l'activité de restriction de SAMHD1, ce mutant tronqué peut être considéré comme constitutivement actif. J'ai ensuite induit l'expression de SAMHD1 dans ces lignées et mesuré leurs vitesses de prolifération pendant plusieurs semaines (Figure 5A). L'expression des différentes constructions a été vérifiée au long de l'expérience (Supplementary Figure 7). La surexpression de SAMHD1 sauvage réduit la vitesse de prolifération des cellules, tandis que la surexpression du mutant inactif ne modifie pas leur vitesse de prolifération. A l'inverse, la surexpression du mutant constitutivement actif ralentit encore plus la vitesse de prolifération des cellules que celle de la protéine sauvage. Cette expérience confirme l'influence de SAMHD1 sur le cycle cellulaire puisque la surexpression de SAMHD1 ralentit la prolifération de façon dépendante de son activité catalytique.

Sachant que SAMHD1 est capable de dégrader les dNTP et que ces mêmes dNTP sont utilisés lors de la réparation des dommages à l'ADN, j'ai voulu savoir si la surexpression de SAMHD1 modifiait la réponse des cellules à différents types de dommages à l'ADN.

J'ai pour cela utilisé des cellules HeLa surexprimant stablement la protéine SAMHD1 sauvage et des cellules contrôle. Celles-ci ont été traitées avec des concentrations croissantes de différents agents causant des dommages à l'ADN pendant trois jours. Le taux de survie des cellules à l'issue du traitement a ensuite été mesuré (Figure 5B). L'étoposide et la camptothécine (CPT) induisent des ruptures double brin de l'ADN en inhibant respectivement les topoisomérases II et I. Le traitement de cellules HeLa par ces drogues induit une mortalité supérieure en présence de SAMHD1. De la même façon, le traitement par la mitomycine C (MMC), qui induit des ruptures double brin et des alkylations (ajouts de groupements hydrocarbonés) de l'ADN provoque une mortalité plus importante des cellules surexprimant SAMHD1. En revanche, le méthanesulfonate de méthyle (MMS), un agent méthylant de l'ADN induit une mortalité comparable des cellules surexprimant SAMHD1 et des cellules contrôles. Ces résultats montrent que la surexpression de SAMHD1 rend les cellules plus sensibles aux ruptures double brin de l'ADN, en favorisant la mort cellulaire par rapport à la réparation de ce type de dommage. Par contre SAMHD1 ne modifie pas la mortalité cellulaire après méthylation de l'ADN.

SAMHD1 module le cycle cellulaire et est régulée en fonction du cycle cellulaire. La découverte de l'influence de SAMHD1 sur la réponse aux lésions de l'ADN m'a donc amenée à me demander si cette protéine est régulée de façon spécifique après induction de ruptures double brin de l'ADN.

J'ai dans un premier temps voulu savoir si les partenaires de SAMHD1 sont modifiés en cas de de rupture double brin de l'ADN. J'ai pour cela traité des cellules HeLa surexprimant SAMHD1 sauvage avec 5nM d'étoposide pendant des durées croissantes. J'ai ensuite purifié les extraits protéiques nucléaires des cellules correspondantes, que j'ai séparés sur des gradients de glycérol. Cette technique permet de séparer les complexes protéiques en fonction de leur taille. Enfin, j'ai observé l'expression de SAMHD1 dans les différentes fractions des gradients par immunoblot (Figure 5C). J'ai ainsi pu montrer que SAMHD1 est normalement localisé dans des fractions intermédiaires correspondant à un complexe moléculaire donné et que le traitement par l'étoposide induit le recrutement de SAMHD1 dans un complexe de plus faible poids moléculaire dès 4 heures de traitement. Ce complexe a une existence transitoire puisque, dès 18 heures après le traitement étoposide, il n'est plus observable sur le gradient de glycérol. Ce complexe de faible poids moléculaire est induit spécifiquement en réponse aux ruptures double brin de l'ADN puisqu'il n'est pas observé après traitement des cellules HeLa par du méthanesulfonate de méthyle (Résultat additionnel 3).



Résultat additionnel 3 : SAMHD1 n'est pas recruté dans un nouveau complexe après méthylation de l'ADN. Des cellules HeLa surexprimant SAMHD1 ont été traitées avec du MMS à 10µg/mL pendant le temps indiqué. Puis les extraits protéiques nucléaires correspondants ont été purifiés et séparés sur gradient de glycérol. L'expression de SAMHD1 dans les différentes fractions a été analysée par immunoblot.

Je me suis ensuite demandé quels étaient les partenaires de SAMHD1 recrutés en cas de rupture double brin de l'ADN. J'ai donc dans un premier temps réalisé des immunoprécipitations de SAMHD1 sur des extraits protéiques de cellules HeLa surexprimant SAMHD1 traitées avec 5nM d'étoposide pendant des durées croissantes, et j'ai regardé si ce traitement étoposide modifiait les interactions de SAMHD1 avec certains de ses partenaires connus (Figure 5D). Si la quantité globale de SAMHD1 n'est pas modifiée après induction de ruptures double brin de l'ADN, son interaction avec la

cycline A est renforcée dès 4 heures après le traitement, c'est-à-dire au moment où SAMHD1 est localisé dans un complexe spécifique. Cette augmentation de l'interaction entre SAMHD1 et la cycline A est spécifique puisque l'interaction entre SAMHD1 et DCAF1 n'est pas renforcée par le traitement étoposide, au contraire.

Sachant que la cycline A, associée à CDK1 ou CDK2, peut phosphoryler SAMHD1 sur la T592, je me suis demandé si l'induction de ruptures double brin de l'ADN favorisait la phosphorylation de SAMHD1. J'ai donc regardé l'expression de SAMHD1 total et de SAMHD1 phosphorylé sur la T592 (pSAMHD1) par immunoblot dans des cellules HeLa traitées avec de l'étoposide. Si, comme précédemment observé, la quantité totale de SAMHD1 ne varie pas, l'expression de pSAMHD1 augmente dès 4 heures après le traitement étoposide et cette phosphorylation est stable puisqu'elle est toujours observable 18 heures après traitement.

L'ensemble de ces données montre que SAMHD1 est spécifiquement régulée en cas de ruptures double brin de l'ADN, notamment par phosphorylation sur la T592. SAMHD1 est donc bien une protéine de réponse aux dommages à l'ADN. Des données de microscopies obtenues par nos collaborateurs confirment d'ailleurs que SAMHD1 localise directement aux sites de ruptures double brin (Figure 5F).

J'ai ainsi montré que SAMHD1 influence le cycle cellulaire mais aussi la réponse à certains dommages à l'ADN. Or le maintien de l'intégrité du génome et le contrôle de la prolifération cellulaire font partie des processus dont le dérèglement peut causer le développement de cancers (Hanahan and Weinberg, 2011). Cela m'a amenée à me demander si l'expression de SAMHD1, observée dans de nombreux types cellulaires, pouvait prévenir la tumorigenèse.

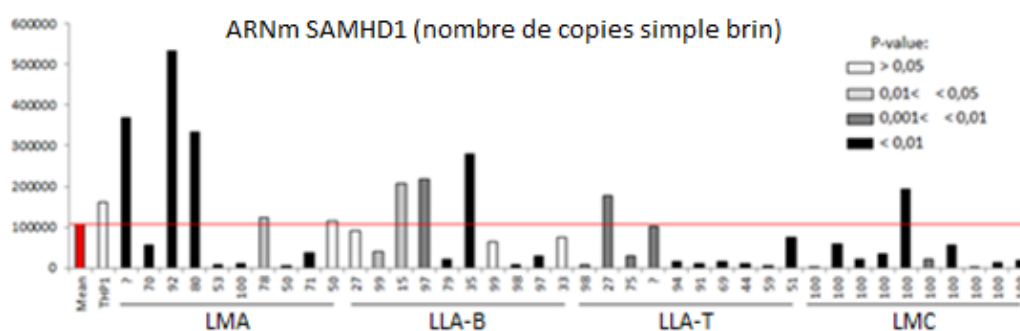
Il est difficile d'étudier le développement de cancers chez des personnes possédant un défaut d'expression de SAMHD1 puisque ce défaut d'expression cause l'AGS et que cette maladie conduit le plus souvent à une mort très précoce. Néanmoins, un des patients AGS connu pour exprimer majoritairement une forme mutée de SAMHD1 (ayant perdu l'exon 15 du gène correspondant) (Rice et al., 2009) a développé une leucémie à l'âge de 24 ans (Figure 1A). Il s'agit d'une leucémie lymphoïde chronique (LLC), qui dans 70% des cas touche des personnes âgées de plus de 50 ans. Son apparition chez un patient de 24 ans peut donc être considérée comme précoce. L'observation de cette corrélation entre perte de SAMHD1 et tumorigenèse précoce, associée aux résultats préalablement obtenus en laboratoire, suggère que SAMHD1 pourrait effectivement jouer un rôle dans la protection contre la tumorigenèse.

C'est dans ce contexte que nous avons choisi de collaborer avec l'équipe du Dr Schuh à Oxford. En effet, leurs analyses ont permis de montrer que, dans deux cohortes indépendantes, 3 à 5% des patients souffrant de LLC présentent des mutations acquises dans le gène SAMHD1 (Figure 2, Table 1 et Supplementary Table 5). Ces mutations sont détectables au moment du diagnostic et, comme celles identifiées dans l'AGS, sont réparties sur l'ensemble du gène (Figure 3A). De plus, leur proportion est plus importante (11%) chez les patients souffrant de tumeurs résistantes au traitement (Table 1).

De notre côté, j'ai montré que, même chez les patients ne présentant pas de mutation dans le gène SAMHD1, le niveau d'expression de SAMHD1 dans les cellules tumorales est inférieur à celui observé dans des PBMC de donneurs sains (Figure 4B et 4C). Ainsi, sur 10 patients étudiés, la quantité d'ARNm SAMHD1 comparée à la moyenne des 7 donneurs sains est significativement réduite ($p < 0,01$) chez 100% des patients. Au niveau protéique, l'expression de SAMHD1 est très faible voire indétectable chez 7 patients sur 10 alors qu'elle semble augmentée chez deux patients. L'existence de patients présentant une quantité réduite d'ARNm SAMHD1 et une augmentation de l'expression de la protéine correspondante suggère que des mécanismes de régulation de l'expression de SAMHD1 restent à élucider. De même, la réduction de l'expression de SAMHD1 chez les patients ne présentant pas de mutation dans le gène correspondant illustre la mise en œuvre de mécanismes de régulation que nous n'avons pas identifiés.

Ces résultats montrent en tous cas que des mutations dans le gène SAMHD1 peuvent être associées au développement de LLC et que, même en l'absence de ces mutations, l'expression de la protéine SAMHD1 est réduite chez une large majorité des patients souffrant de LLC.

J'ai ensuite voulu savoir si la perte de l'expression de SAMHD1 était une spécificité des LLC ou si cette caractéristique pouvait être retrouvée dans d'autres types de cancers. J'ai pour cela analysé l'expression des ARNm SAMHD1 par PCR quantitative chez des patients souffrant de quatre autres types de leucémie (Résultat additionnel 4). Le niveau d'expression chez chacun des 40 patients a été comparé à celui observé chez les sept donneurs sains préalablement étudiés (Figure 4B et C). Ces analyses montrent que l'expression de SAMHD1 est significativement réduite chez plus de 50% des patients souffrant de cancers hématopoïétiques (29 patients sur 40). La fréquence des patients perdant l'expression de SAMHD1 varie selon le cancer considéré (90% dans les leucémies myéloïdes chroniques, 80% dans les leucémies lymphoïdes aiguës T, 40% dans les leucémies lymphoïdes aiguës B et 50% dans les leucémies myéloïdes aiguës) mais la réduction de l'expression de SAMHD1 est retrouvée dans l'ensemble des leucémies étudiées et pas uniquement chez les patients souffrant de LLC.



Résultat additionnel 4 : la quantité d'ARNm SAMHD1 est affectée dans différentes leucémies. L'ARNm SAMHD1 a été quantifié dans 40 échantillons d'ARN issus de prélèvements de patients souffrant de quatre types de leucémies (LMA : leucémie myéloïde aiguë, LLA-B : leucémie lymphoïde aiguë B, LLA-T : leucémie lymphoïde aiguë T, L : leucémie myéloïde chronique). Cette valeur a été comparée aux sept échantillons de donneurs sains utilisés dans l'article (code couleur indiquant les p-values) et dont la valeur moyenne d'expression de SAMHD1 est ici représentée en rouge. Les nombres en abscisse indiquent la proportion de cellules tumorales dans l'échantillon initialement obtenu.

De plus, une collaboration avec le personnel de Xentech a permis de montrer que des diminutions de l'expression de SAMHD1 sont aussi observées chez certaines patientes souffrant de cancers du sein (Supplementary Figure 6D). Ce phénotype n'est donc pas une spécificité des leucémies et peut être retrouvé dans des tumeurs solides.

Le projet principal que j'ai mené pendant ma thèse a permis de montrer que SAMHD1 est faiblement exprimé et que le gène correspondant est fréquemment muté dans les lignées utilisées en laboratoire. J'ai également mis en évidence un ralentissement du cycle cellulaire lors de la surexpression de SAMHD1 et montré que cette protéine favorise la mort cellulaire en cas de rupture double brin de l'ADN. J'ai d'ailleurs montré que la protéine SAMHD1 elle-même est régulée en cas de rupture double brin de l'ADN, notamment par phosphorylation sur la T592. Enfin, des travaux menés en collaboration avec d'autres équipes ont permis de montrer que la perte de SAMHD1, due à la mutation du gène correspondant ou à d'autre(s) mécanisme(s) non élucidé(s), est associée au développement de différentes tumeurs, aussi bien leucémiques que solides.

Ces résultats montrent que SAMHD1 est un acteur clé de l'homéostasie cellulaire et que sa régulation est essentielle.

Discussion

Implications thérapeutiques nouvelles de SAMHD1

SAMHD1 et résistances à certaines chimiothérapies

Des expériences réalisées dans des lignées cellulaires au laboratoire m'ont permis de montrer que, dans des cellules confrontées à des ruptures double brin de l'ADN, la surexpression de SAMHD1 augmente la mortalité trois jours après traitement. Ces résultats ont été observés dans une lignée issue d'un cancer primaire et exprimant très faiblement SAMHD1.

La surexpression d'un gène à l'expression réduite dans les cellules tumorales a donc permis de rendre celles-ci plus sensibles aux ruptures double brin de l'ADN. Or les produits utilisés dans cette expérience (étoposide et mitomycine C) ou leurs dérivés (camptothécine) sont des agents utilisés dans le traitement des cancers par chimiothérapie. Ces résultats, transposés dans le contexte du traitement d'un cancer primaire, suggèrent que la perte de SAMHD1 au cours du développement cancéreux peut entraîner la perte de la sensibilité à certains traitements. Autrement dit que la perte de SAMHD1 pourrait être à l'origine de résistances à la chimiothérapie (Figure 32).

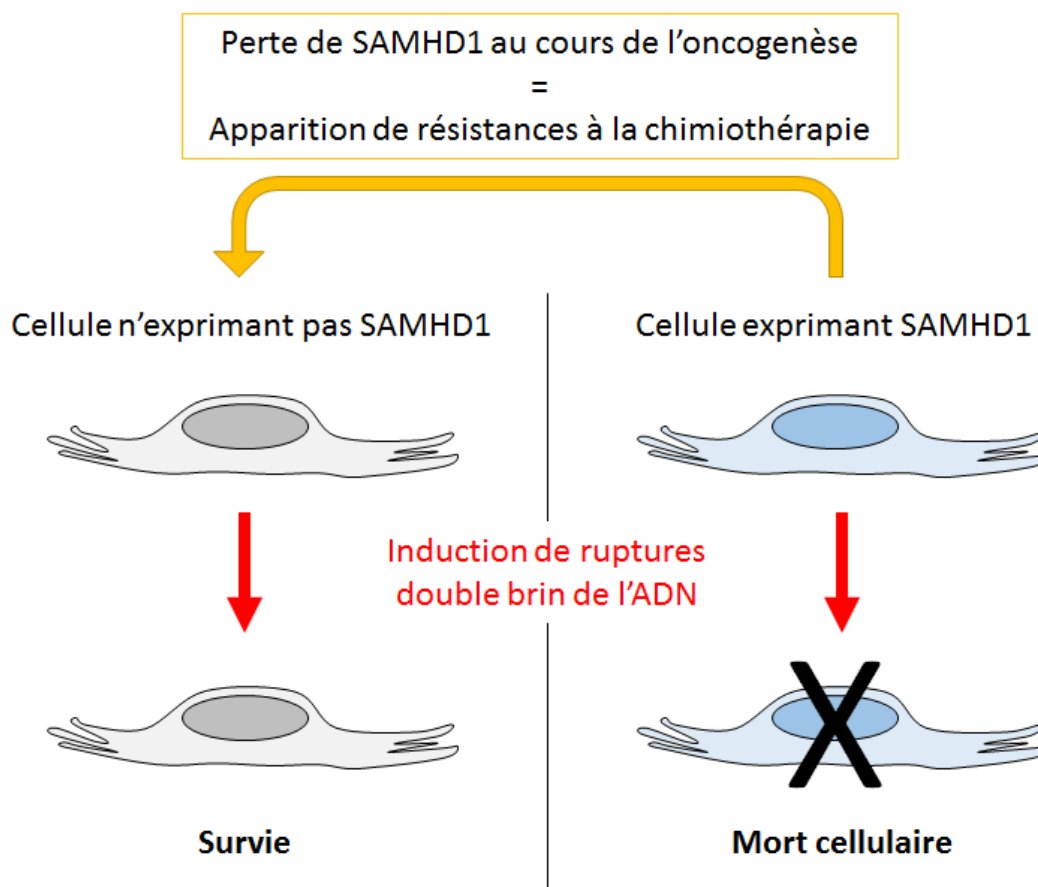


Figure 32 : schéma du principe via lequel la perte de SAMHD1 peut être à l'origine de résistances à certaines chimiothérapies.

La fréquence des mutations du gène SAMHD1 est plus de deux fois plus élevée dans les LLC résistantes au traitement que dans l'ensemble des LLC au moment du diagnostic (11% contre 3 à 5 %). Cette observation va dans le sens d'un rôle de SAMHD1 dans la sensibilité à la chimiothérapie. Néanmoins des informations plus détaillées concernant les types de traitements utilisés pour traiter les patients souffrant de tumeurs résistantes seraient nécessaires pour prolonger cette réflexion. Il

serait notamment intéressant de voir si les traitements par chimiothérapie induisant des ruptures double brin de l'ADN conduisent à la sélection de cellules tumorales n'exprimant plus SAMHD1.

Des analyses de l'expression de SAMHD1 dans des cohortes de patients souffrant de cancers résistants à la chimiothérapie seraient extrêmement intéressantes et pourraient aboutir à une adaptation des traitements utilisés. En effet, nos données suggèrent que la perte de SAMHD1 pourrait être un facteur de résistance aux ruptures double brin de l'ADN. Si cette hypothèse est juste, le traitement de tumeurs comprenant des cellules n'exprimant plus SAMHD1 par des agents induisant des ruptures double brin de l'ADN pourrait avoir un effet néfaste sur l'évolution de la maladie. Ce traitement aboutirait à la sélection des cellules n'exprimant plus SAMHD1, qui seraient les plus résistantes à la chimiothérapie et donc les plus difficiles à éliminer par la suite. De plus, ces cellules survivant à un traitement mutagène sont susceptibles d'accumuler de nouvelles mutations pouvant également aggraver le pronostic.

Ainsi, à l'époque de la personnalisation des thérapies anticancéreuses, l'évaluation de l'expression de SAMHD1 par les cellules tumorales pourrait être un élément déterminant dans le choix de l'utilisation ou non de certains agents de chimiothérapie. En effet, il existe des agents de chimiothérapie n'induisant pas de rupture double brin de l'ADN. Parmi ceux-ci, on trouve des agents alkylants qui, d'après mes résultats, agissent de la même façon quel que soit le niveau d'expression de SAMHD1. La prise en compte du niveau d'expression de SAMHD1 dans les tumeurs permettrait donc d'optimiser les traitements par chimiothérapie.

Des études plus poussées en laboratoire pourraient permettre de voir quels agents utilisés en chimiothérapie ont un effet évoluant en fonction du niveau d'expression de SAMHD1. La réponse à ceux dont l'efficacité diminue en cas de perte de SAMHD1 pourrait alors être évaluée dans des cohortes existantes, en fonction de l'expression de SAMHD1 par les cellules tumorales des différents patients traités. Cela permettrait de déterminer si la perte de SAMHD1 au cours du développement tumoral est réellement à l'origine de résistances à certains traitements et dans quels types de cancers. Ces informations primordiales pourraient alors permettre d'adapter au mieux les traitements utilisés en fonction du profil d'expression des tumeurs et de son évolution au cours de la thérapie.

Perte de SAMHD1 et tumorigenèse

Nos travaux montrent que, dans le cas de différentes leucémies et des cancers du sein, il y a une corrélation entre la perte de l'expression de SAMHD1 et le processus d'oncogenèse. Les données que j'ai obtenues sur des lignées cellulaires et le caractère précoce des mutations du gène SAMHD1 suggèrent par ailleurs que la perte de SAMHD1 pourrait être un des moteurs du développement cancéreux.

Le fait que SAMHD1 soit exprimé dans de nombreux tissus humains (Schmidt et al., 2015) indépendamment de leur sensibilité à l'infection par le VIH renforce le point de vue selon lequel, avant d'être un facteur de restriction, SAMHD1 est un gardien de l'intégrité cellulaire.

De plus, depuis la publication de notre article, une seconde étude a mis en évidence une corrélation entre la perte de SAMHD1 et le développement cancéreux (Wang et al., 2014a). Cette étude a été réalisée sur des adénocarcinomes pulmonaires dont le caractère solide a permis aux auteurs de comparer l'expression de SAMHD1 entre le tissu cancéreux et le tissu sain chez cinq patients. Comme nous, ils ont observé une diminution de l'expression de SAMHD1 au niveau de la protéine et de l'ARNm. Ils n'ont pas séquencé le gène correspondant mais, au-delà des variations quantitatives, la détection de protéine SAMHD1 de taille identique dans les prélèvements cancéreux et dans les prélèvements sains suggère que le gène correspondant est inhibé mais non muté. Ils ont montré que cette inhibition est due à la méthylation du promoteur du gène codant pour la protéine

SAMHD1 et que ce même mécanisme réduit l'expression de SAMHD1 dans deux lignées d'adénocarcinomes pulmonaires.

De mon côté, j'ai montré que chez les patients souffrant de LLC, l'expression de SAMHD1 peut être réduite en l'absence de mutation dans le gène correspondant. Il serait intéressant de vérifier l'état de méthylation du promoteur de SAMHD1 chez ces patients pour savoir si, comme dans le cas du cancer du poumon, l'expression de SAMHD1 est réprimée par ce mécanisme. De plus, dans deux des échantillons que j'ai étudiés, les ARNm SAMHD1 étaient faiblement exprimés sans que le niveau d'expression de la protéine correspondante ne soit affecté. Cela suggère qu'il doit exister un mécanisme de régulation post-transcriptionnel de l'expression de SAMHD1, qui reste encore à élucider.

Globalement, quels que soient les mécanismes mis en jeu, notre étude a montré un lien entre diminution de l'expression de SAMHD1 et tumorigenèse. Nous avons mis ce lien en évidence dans cinq types de leucémie et dans un cancer solide : le cancer du sein. Depuis, une autre équipe a confirmé cette observation dans un second type de cancer solide : le cancer du poumon. Des études à plus large échelle comprenant des types de cancers plus nombreux seraient nécessaires pour évaluer l'impact de la perte de SAMHD1 dans le développement cancéreux en général et mettre en évidence d'éventuelles spécificités.

Les résultats obtenus sur les cancers pulmonaires vont dans le même sens que les nôtres et soulignent le lien entre diminution de l'expression de SAMHD1 et développement cancéreux en s'appuyant sur un autre type de cancers que ceux que nous avons étudiés. Au-delà de la caractérisation de la particularité (ou non) de ce phénotype à certains types de cancers, d'autres études sont nécessaires pour mettre en évidence un lien de causalité formel entre déplétion de SAMHD1 et cancer *in vivo*. L'ensemble des données accumulées à ce jour suggère qu'un tel lien existe, mais aucune preuve n'a réellement été fournie.

Le cas du patient AGS ayant développé une LLC précoce ne peut être considéré comme démonstratif d'un lien de causalité. En effet, ce patient possède également une mutation dans le gène ADAR1 et l'inflammation chronique liée à sa maladie génétique peut constituer un facteur de risque vis-à-vis de l'oncogenèse. De même, si les patients souffrant de LLC chez lesquels nous avons identifié des mutations acquises dans le gène SAMHD1 ne présentent pour la plupart pas de mutation dans les gènes connus comme origines possibles de LLC, rien ne permet d'exclure qu'une mutation non étudiée dans un autre gène soit la cause première de la maladie.

Il serait notamment intéressant d'étudier les lignées murines n'exprimant plus SAMHD1 qui ont été générées récemment (Behrendt et al., 2013; Rehwinkel et al., 2013) pour voir si la perte de cette protéine rend les animaux plus prompts à développer des cancers, en l'absence de traitement et après traitement avec différents agents oncogènes. La mise en place de modèles de tumorigenèse chez ces souris et chez des souris contrôles pourrait également permettre d'étudier la réponse de tumeurs n'exprimant pas SAMHD1 à différents traitements de chimiothérapie.

S'il était clairement démontré que la perte de SAMHD1 peut être un des premiers événements favorisant le développement cancéreux, l'analyse de l'expression de cette protéine pourrait être utilisée comme marqueur dans l'étude de biopsies ou de prélèvements sanguins à visées diagnostiques.

Mise en évidence de fonctions cellulaires de SAMHD1

SAMHD1 et progression du cycle cellulaire

Plusieurs équipes ont étudié l'effet de la réduction de l'expression de SAMHD1 sur le cycle cellulaire. Toutes ont montré que la perte de SAMHD1 induit un ralentissement du cycle cellulaire associé à une augmentation de la proportion des cellules en phase G1 (Franzolin et al., 2013; Kretschmer et al., 2014). Cette perturbation du cycle cellulaire est associée à l'augmentation de la quantité des différents dNTP et à la dérégulation de leurs proportions respectives. Ces résultats

montrent qu'une perturbation de la quantité de SAMHD1 entraîne un dysfonctionnement du cycle cellulaire, lui-même vraisemblablement lié à la modification de la quantité des dNTP intracellulaires. Ainsi, en l'absence de SAMHD1, l'accumulation anormale de dNTP en dehors de la phase S favoriserait un arrêt du cycle cellulaire en fin de phase G1.

De mon côté, j'ai montré que la surexpression de SAMHD1 induit également un ralentissement de la prolifération cellulaire et que ce phénotype dépend de la présence d'un domaine HD fonctionnel. Cela suggère que la surexpression de SAMHD1 peut également perturber le cycle cellulaire. J'ai fait l'hypothèse que cela était dû à une diminution de la quantité de dNTP disponibles et donc à un ralentissement de la phase S, mais ne l'ai pas vérifié expérimentalement. Des résultats comparables ont depuis été obtenus avec une lignée cellulaire issue d'une tumeur pulmonaire (Wang et al., 2014a). La surexpression de SAMHD1 dans ces cellules ralentit leur prolifération et est associée à une diminution de près de 40% de la quantité de dNTP. Cela renforce l'hypothèse selon laquelle SAMHD1 ralentit le cycle cellulaire en diminuant la quantité de dNTP disponibles (Figure 33).

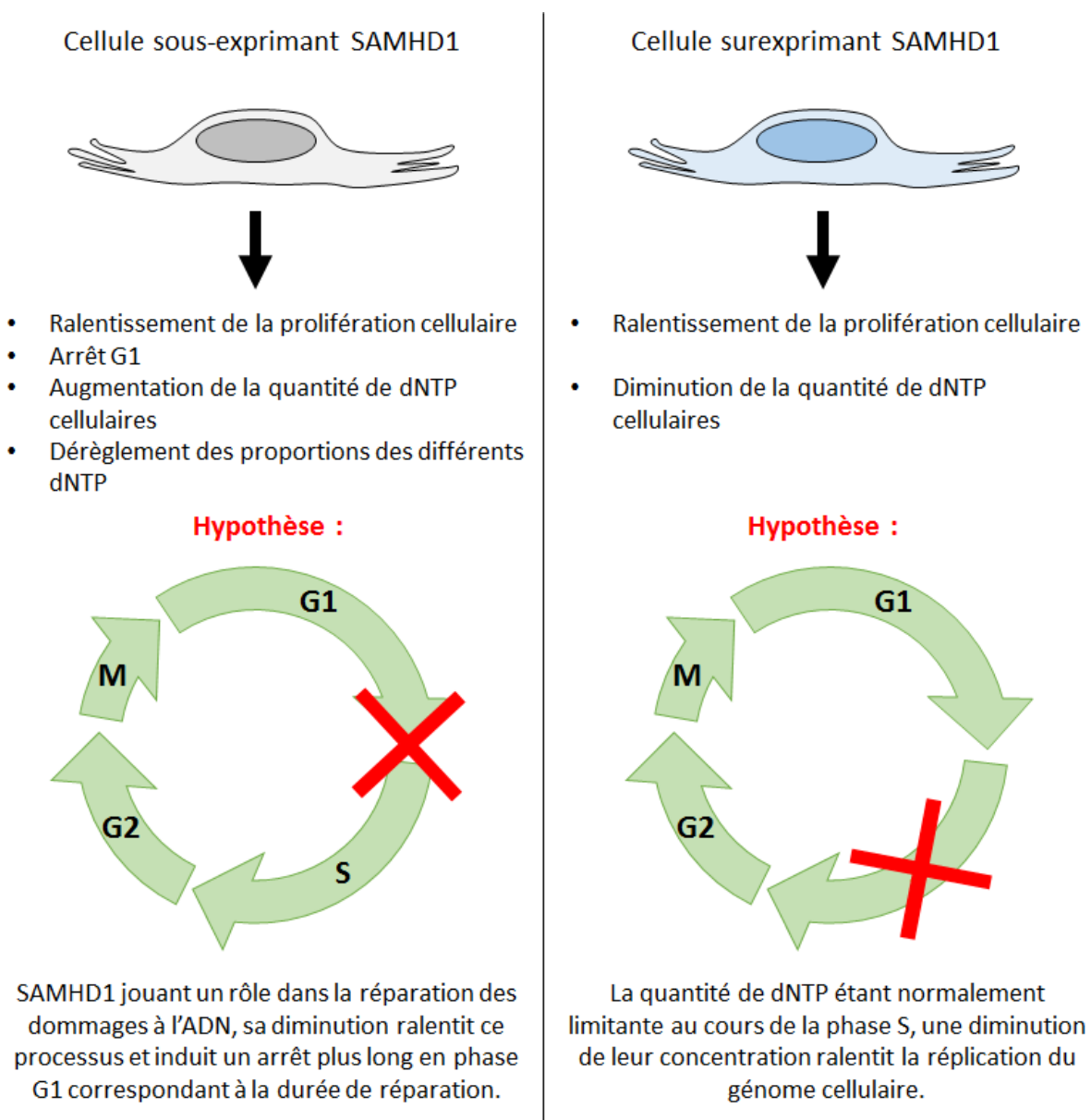


Figure 33 : schéma récapitulant les effets observés de SAMHD1 sur le cycle cellulaire et les hypothèses proposées pour les expliquer.

L'ensemble de ces résultats montre que la surexpression et la sous expression de SAMHD1 ralentissent toutes deux le cycle cellulaire, vraisemblablement en raison de la perturbation de la quantité de dNTP cellulaires qui en découle. En effet, il est connu que la quantité de dNTP est régulée très finement au cours du cycle cellulaire et reste limitante même pendant la phase S (Poli et al., 2012), ce qui empêche l'apparition de mutations dues à un excès de dNTP (Reichard, 1988). On peut ainsi imaginer que la diminution de cette quantité de dNTP en cas de surexpression de SAMHD1 entrave la progression de cette phase S au cours de laquelle les dNTP sont utilisés pour la réplication du génome cellulaire.

Parallèlement, la sous-expression de SAMHD1 pourrait perturber le bon fonctionnement des autres processus utilisant des dNTP, comme la réplication du génome mitochondrial (Leshinsky-Silver et al., 2011) ou la réparation des dommages à l'ADN. Ces dysfonctionnements pourraient être à l'origine de l'arrêt G1/S observé en cas de sous expression de SAMHD1. Il est notamment connu que le passage en phase S n'est pas possible en présence de lésions à l'ADN non réparées (Boye et al., 2009). Or j'ai pu montrer que SAMHD1 est régulée en cas de lésion à l'ADN et recrutée aux sites de ruptures double brin. On peut ainsi imaginer que l'absence de SAMHD1 perturbe la réparation de ce type de lésions, ce qui expliquerait les perturbations de la transition G1/S observées par d'autres équipes en l'absence de SAMHD1.

Ainsi, même si les preuves expérimentales sont pour l'instant manquantes on peut supposer que la surexpression et la sous expression de SAMHD1 perturbent le cycle cellulaire via des mécanismes différents, la première diminuant la vitesse de la phase S et la seconde induisant une accumulation de lésions à l'ADN bloquant le cycle cellulaire en phase G1 (Figure 33). Cela expliquerait pourquoi l'augmentation et la diminution de la quantité de SAMHD1 dans les cellules ont le même effet sur le cycle cellulaire au lieu d'avoir des effets opposés.

SAMHD1 et réparation des dommages à l'ADN

Les résultats que j'ai obtenus pendant ma thèse ont été les premiers à mettre en évidence un rôle de SAMHD1 dans la réponse aux dommages à l'ADN et plus particulièrement aux ruptures double brin. J'ai montré qu'en cas de traitement avec des agents chimiques induisant des ruptures double brin de l'ADN, la surexpression de SAMHD1 favorise la mort cellulaire. En revanche elle n'a aucun effet sur la survie des cellules après traitement avec des agents alkylant l'ADN.

Depuis, d'autres données obtenues à partir de l'étude de fibroblastes de patients AGS ont montré que la perte de l'expression de SAMHD1 favorise l'accumulation de lésions à l'ADN. De plus, si ces cellules de patients AGS ne présentent pas plus de ruptures double brin que des cellules équivalentes de donneurs sains exprimant normalement SAMHD1, elles présentent en revanche plus de ruptures double brin que les cellules contrôles après traitement avec des faibles doses de rayons ultra-violets (20J/m²) (Kretschmer et al., 2014). Ces résultats suggèrent un rôle de SAMHD1 dans l'efficacité de la prise en charge de ces lésions. Notre observation du recrutement de SAMHD1 aux sites de ruptures double brin de l'ADN renforce cette hypothèse et laisse penser que SAMHD1 pourrait agir directement aux sites de clivage.

Du point de vue moléculaire, beaucoup de travail reste à accomplir pour comprendre la façon dont SAMHD1 intervient lors de la réparation des coupures double brin de l'ADN. A l'heure actuelle tout ce que nous savons est que SAMHD1 est recruté au niveau des sites de clivage et qu'il est recruté dans un complexe spécifique après induction de ruptures double brin.

Compte-tenu de la complexité des mécanismes de réparation des coupures double brin de l'ADN, il semble primordial de commencer par répondre à deux questions. D'une part : est-ce que SAMHD1 interagit spécifiquement avec des protéines impliquées dans la réparation des clivages de l'ADN après induction de ces lésions ? Et d'autre part : par quel mécanisme SAMHD1 est-il recruté directement au niveau des lésions ?

En effet, identifier les partenaires de SAMHD1 dans un contexte de dommages à l'ADN permettrait de comprendre à quel niveau du mécanisme de réparation ce nouvel acteur est susceptible d'intervenir, ce qui constitue une étape clé dans la compréhension de la façon dont SAMHD1 module la réparation des dommages à l'ADN. Par ailleurs, cela permettrait de concevoir des expériences pour déterminer comment SAMHD1 est recruté aux sites de ruptures de l'ADN. En effet, on peut imaginer que SAMHD1 y soit recruté via son interaction avec une protéine connue de réparation des dommages à l'ADN, mais aucun élément ne permet actuellement de l'affirmer. SAMHD1 peut tout aussi bien être recruté directement via sa capacité d'interagir avec les acides nucléiques, indépendamment des protéines de réparation connues.

Ainsi, la connaissance des partenaires de SAMHD1 et de son mécanisme de recrutement en cas de ruptures double brin de l'ADN sont les premières étapes qui permettraient à terme de comprendre la façon dont cette protéine intervient dans la réparation des dommages à l'ADN.

A l'échelle cellulaire, plusieurs situations sont possibles en cas de ruptures double brin de l'ADN. Ou bien celles-ci sont réparées, quitte à induire un arrêt temporaire du cycle cellulaire en phase G1, et la cellule continue de se diviser normalement, ou bien elles ne sont pas réparées et finissent par induire la mort cellulaire, ou bien elles ne sont pas réparées ou mal réparées et peuvent conduire à la survie d'une cellule accumulant des mutations et dont l'instabilité génétique peut enclencher le processus de tumorigenèse..

A cette échelle, les résultats actuellement publiés mettent clairement en valeur le rôle de SAMHD1. En effet, en l'absence de cette protéine, les ruptures double brin de l'ADN s'accumulent alors qu'en excès de cette protéine les cellules sont plus facilement tuées par des agents chimiques provoquant des ruptures double brin de l'ADN. SAMHD1 est donc un des éléments impliqués dans l'orientation de la cellule vers la mort cellulaire plutôt que vers la réparation quand les dommages causés à l'ADN sont trop importants pour être efficacement réparés.

Dans cette optique, SAMHD1 peut être considéré comme une protection contre l'instabilité génomique et donc contre la tumorigenèse. Cette analyse de la fonction cellulaire de SAMHD1 est cohérente avec les observations que nous avons pu faire sur des lignées transformées et des cellules cancéreuses primaires : une proportion non négligeable de celles-ci perd l'expression de SAMHD1.

Fonctions cellulaires et activités catalytiques de SAMHD1

A ce jour, trois fonctions cellulaires de SAMHD1 ont été caractérisées, en partie par mes travaux de doctorat. J'ai montré que SAMHD1 module la réparation de certains dommages à l'ADN ainsi que la progression du cycle cellulaire. Il était également connu que SAMHD1 est un régulateur négatif de l'immunité innée. Quelle que soit la fonction cellulaire considérée, aucun mécanisme clair n'a été identifié.

Comme rappelé en introduction du présent manuscrit, deux activités catalytiques de SAMHD1 ont été caractérisées à ce jour : une activité dNTP triphosphohydrolase et une activité nucléase. Si les substrats et les mécanismes de régulation allostérique de la première ont été précisément caractérisés, la seconde reste encore assez mal connue et les résultats contradictoires obtenus jusqu'à aujourd'hui ne permettent pas d'avoir une idée précise des substrats concernés.

Un des angles d'approche pour élucider les mécanismes à l'origine des fonctions cellulaires de SAMHD1 pourrait être de caractériser la ou les activité(s) catalytique(s) de cette protéine impliquée(s) dans les différentes fonctions. En effet, aucun élément ne permet d'affirmer à l'heure actuelle que l'une ou l'autre des fonctions de SAMHD1 est due à son activité dNTPase ou à son activité nucléase.

Cela découle notamment du fait que ces deux activités catalytiques sont portées par les mêmes domaines de la protéine et qu'aucune construction ne permettait de les discriminer. Seule une publication fait état de mutants de SAMHD1 ne possédant qu'une seule des deux activités catalytiques (Ryoo et al., 2014) : le mutant D137N a une activité RNase comparable à celle de SAMHD1 sauvage

mais aucune activité dNTPase, à l'inverse du mutant Q548N. Ces résultats, datant d'environ un an, n'ont pas été reproduits depuis et aucune donnée concernant les fonctions cellulaires de SAMHD1 n'a été publiée en utilisant ces mutants.

La principale caractéristique qui a été utilisée pour essayer de séparer le rôle des deux activités de SAMHD1 est le phénotype associé à la phosphorylation sur la T592. En effet, cette phosphorylation inhibe l'effet de restriction de SAMHD1 sur le VIH-1 mais pas son activité dNTPase. Cela suggère qu'une autre activité catalytique de SAMHD1 est inhibée par cette phosphorylation. Néanmoins, si cette phosphorylation stabilise le génome du VIH-1 in vivo, elle ne modifie pas l'activité RNase de SAMHD1 in vitro, ni sa capacité à interagir avec les acides nucléiques (Ryoo et al., 2014). L'utilisation de mutants phospho-mimétiques ou non phosphorylables de SAMHD1 ne permet donc pas clairement de séparer les différentes activités catalytiques de cette protéine. Elle peut néanmoins permettre de conclure qu'un phénomène n'est pas dépendant que de l'activité dNTPase de SAMHD1 s'il est affecté par la phosphorylation T592, puisque celle-ci ne modifie pas l'activité dNTPase de la protéine.

Effet de SAMHD1 sur le cycle cellulaire

Concernant l'effet de SAMHD1 sur le cycle cellulaire, nous sommes les seuls à avoir utilisé des mutants de SAMHD1 susceptibles de nous donner des indications sur l'importance des différentes activités catalytiques de la protéine dans les phénotypes observés.

D'après mes résultats, le ralentissement de la prolifération cellulaire est plus marqué en cas de surexpression d'un mutant de SAMHD1 non phosphorylable sur la T592 qu'en cas de surexpression de la protéine sauvage. Ceci suggère que SAMHD1 sauvage a un effet moindre sur la vitesse de prolifération en raison de sa phosphorylation partielle sur la T592. Or il a été montré que cette phosphorylation n'a pas d'effet sur l'activité dNTPase de SAMHD1. Ces données suggèrent donc que la seule activité dNTPase de SAMHD1 ne suffit pas à expliquer son effet sur le ralentissement du cycle cellulaire. Cette conclusion doit néanmoins être considérée avec prudence. En effet, j'ai utilisé un mutant tronqué pour l'ensemble du domaine C-terminal de la protéine et non un mutant ponctuel de la T592. Il est possible que cette délétion affecte l'activité de SAMHD1 via d'autres mécanismes que la perte d'un site phosphorylable.

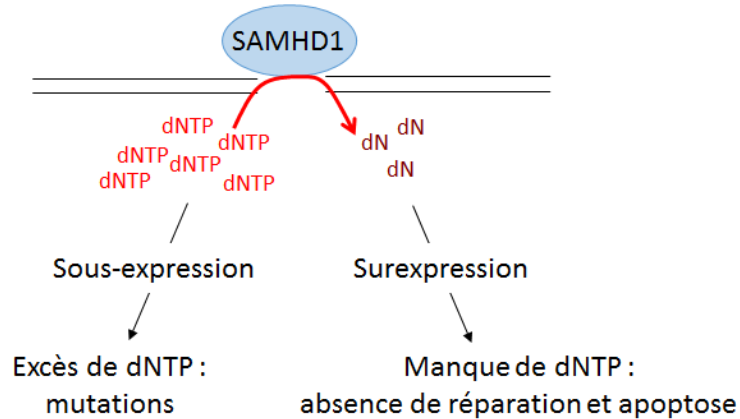
Ainsi, si on considère que le mutant tronqué au niveau du domaine C-terminal de SAMHD1 n'affecte l'activité de la protéine que via son absence de phosphorylation sur la T592, mes résultats montrent que le ralentissement de la prolifération cellulaire en l'absence de SAMHD1 n'est pas uniquement dû à son activité dNTPase. Il n'est cependant pas exclu que cette activité soit responsable d'une partie du phénotype. Cela pourrait suggérer que, contrairement à notre hypothèse de départ, SAMHD1 ne ralentit pas le cycle cellulaire qu'en ralentissant la réplication du génome lors de la phase S. La reproduction de l'expérience de prolifération cellulaire en utilisant des mutants ponctuels de SAMHD1 (non phosphorylable, phospho-mimétique et séparant les deux activités catalytiques connues) permettrait d'établir des hypothèses plus précises sur l'origine du ralentissement de la prolifération cellulaire en cas de surexpression de SAMHD1.

En ce qui concerne l'absence de SAMHD1 et l'arrêt G1/S du cycle cellulaire, l'étude de l'importance des différentes activités est plus complexe. En effet, s'il est possible d'envisager des expériences de complémentations avec des mutants séparant les deux activités catalytiques connues, il est impératif dans ce contexte de mettre au point un protocole permettant des niveaux d'expression équivalents des différents variants de SAMHD1 utilisés.

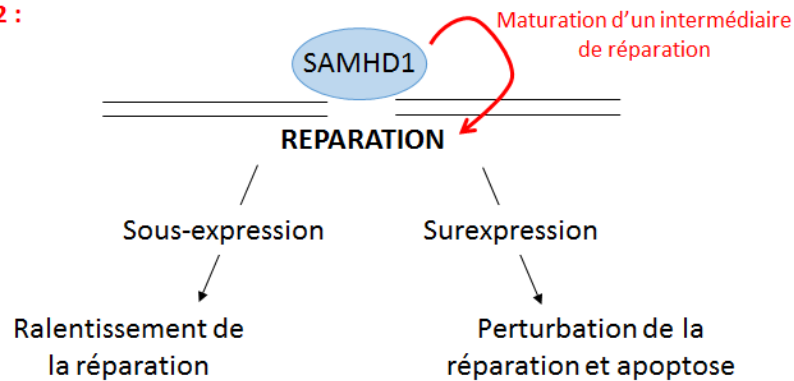
SAMHD1 et réparation des dommages à l'ADN

A propos des effets de SAMHD1 sur la réponse aux dommages à l'ADN, aucune étude n'a été réalisée avec des mutants permettant de conclure sur les activités catalytiques impliquées. Plusieurs hypothèses peuvent cependant être faites a priori (Figure 34).

Hypothèse 1 :



Hypothèse 2 :



Hypothèse 3 :

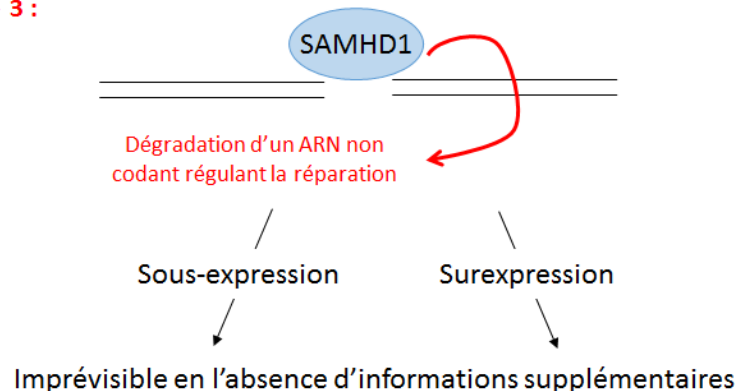


Figure 34 : schéma de trois hypothèses pouvant expliquer la façon dont SAMHD1 intervient dans la réponse aux dommages à l'ADN. Hypothèse 1 : SAMHD1 agit en diminuant localement la quantité de dNTP disponibles. Hypothèse 2 : SAMHD1 permet la maturation de structures intermédiaires de la réparation. Hypothèse 3 : SAMHD1 dégrade un ARN non codant régulant la réparation.

SAMHD1 pourrait agir en diminuant localement la quantité de dNTP présents au niveau des sites de clivage pour favoriser une réparation fiable des lésions à l'ADN. Dans cette hypothèse,

l'absence de SAMHD1 favoriserait l'accumulation d'erreurs lors de la réparation, susceptibles de conduire à une instabilité génétique globale. A l'inverse, la surexpression de SAMHD1 diminuerait la quantité de dNTP disponibles au niveau des sites de clivage au point d'empêcher une réparation efficace, ce qui induirait la mort cellulaire par accumulation de lésions non réparées. Cette hypothèse est cependant en contradiction avec celle supposant que la sous-expression de SAMHD1 induit un arrêt en phase G1 lors du cycle cellulaire en raison de l'accumulation de lésions à l'ADN non réparées. En effet, si SAMHD1 agissait en diminuant la quantité de dNTP disponibles au niveau des sites de clivage, la diminution de l'expression de SAMHD1 favoriserait la réparation de ces lésions, mais avec accumulation de mutations.

On peut également supposer que SAMHD1 agirait via son activité nucléase en clivant directement des structures intermédiaires de la réparation. Dans ce cas, l'augmentation de la concentration de SAMHD1 pourrait nuire à la réparation des lésions en provoquant des clivages prématurés mal pris en charge par le reste de la machinerie de réparation. Cette configuration pourrait favoriser la mort cellulaire en cas de lésions à l'ADN dans un contexte de surexpression de SAMHD1, ce qui correspond au phénotype que j'ai pu observer. A l'inverse la diminution de l'expression de SAMHD1 ralentirait le processus de réparation, ce qui expliquerait l'arrêt G1/S observé dans cette configuration.

Par ailleurs de plus en plus de données montrent que certains ARN non codants jouent des rôles essentiels dans la réponse aux dommages à l'ADN et dans sa régulation (Liu and Lu, 2012).

Les microARN sont notamment impliqués dans cette réponse. Il a ainsi été montré que la suppression d'effecteurs généralistes de la voie des microARN comme Dicer et Ago2 induit une augmentation de la mortalité après traitement UV et abolit l'arrêt G1 normalement induit en cas de lésion à l'ADN (Pothof et al., 2009). Ces microARN impliqués dans la réparation des dommages à l'ADN peuvent intervenir à différents niveaux de la voie en inhibant la traduction de protéines comme H2AX ou p53 (Liu and Lu, 2012). Ils peuvent également être induits par certaines protéines de réponse aux dommages à l'ADN : p53 est ainsi capable d'induire l'expression de microARN inhibant Mdm2. Mdm2 induisant la dégradation de p53 par le protéasome, p53 permet ainsi l'inhibition d'un de ses propres inhibiteurs (Liu and Lu, 2012).

Différents longs ARN non codants (lncRNA) ont également été caractérisés comme modulant la réponse aux dommages à l'ADN. Il a notamment été montré que le lncRNA-p21 est induit par p53 et joue un rôle dans l'induction de l'apoptose après dommages à l'ADN (Huarte et al., 2010). A l'inverse, le lncRNA PANDA, dont l'expression est également induite par p53 en cas de lésions à l'ADN, limite cette apoptose (Hung et al., 2011).

Or il a été montré par plusieurs équipes que SAMHD1 peut dégrader les ARN simple brin, au moins in vitro (Beloglazova et al., 2013; Ryoo et al., 2014). On pourrait donc imaginer que SAMHD1 module la réparation des dommages à l'ADN en dégradant des ARN non codants qui sont eux-mêmes impliqués dans la régulation de ce phénomène. Compte-tenu de la diversité des ARN non codants impliqués dans la réparation des dommages à l'ADN et de leurs mécanismes d'action, il est difficile d'analyser les effets de SAMHD1 dans le contexte de cette hypothèse.

En l'état actuel des connaissances, il serait intéressant de comparer les niveaux d'expression de différents ARN non codants impliqués dans la réparation dans des cellules exprimant ou non SAMHD1 après traitement par des agents endommageant l'ADN. Cela permettrait de voir si la présence de SAMHD1 corrèle avec la diminution de l'expression d'un ou plusieurs de ces ARN non codants. L'identification d'ARN non codants correspondant à ce critère pourrait alors servir de point de départ à un projet de recherche visant à étudier l'effet de SAMHD1 sur ces ARN dans le contexte de la réparation des dommages à l'ADN.

Aucune de ces hypothèses n'est en contradiction avec mes observations d'un effet de SAMHD1 sur la survie cellulaire en cas de rupture double brin de l'ADN mais pas en cas de méthylation. De même, aucune n'est incompatible avec le recrutement de SAMHD1 directement au niveau des sites

de rupture double brin de l'ADN. Il est donc impossible à l'heure actuelle d'affirmer qu'une des deux activités catalytiques identifiées de SAMHD1 est à l'origine de sa fonction dans la réponse aux dommages à l'ADN.

La seule piste de réflexion concrète actuellement disponible découle des informations connues sur la régulation de SAMHD1 en cas de dommages à l'ADN. En effet, j'ai montré que le traitement avec un agent causant des ruptures double brin de l'ADN induit la phosphorylation de SAMHD1 sur la T592. Or cette phosphorylation n'a pas d'effet sur l'activité dNTPase de cette protéine. Cela ne permet cependant pas d'affirmer que cette activité n'est pas requise pour le rôle de SAMHD1 dans la réponse aux dommages à l'ADN. En effet, on pourrait au contraire imaginer que cette phosphorylation permet d'inhiber d'autres activités catalytiques de SAMHD1 et maximise la proportion de SAMHD1 cellulaire actif en tant que dNTPase. Une fois de plus, la surexpression des mutants permettant de séparer les deux activités catalytiques de SAMHD1 semble être le moyen le plus efficace de distinguer leurs rôles respectifs dans la fonction de SAMHD1 en tant que protéine de réponse aux dommages à l'ADN.

SAMHD1 et réponse immunitaire innée

Le fait que des mutations innées dans le gène SAMHD1 induisent le développement de l'AGS, un syndrome caractérisé par une forte réponse inflammatoire (Rice et al., 2009), montrent que la protéine correspondante joue un rôle dans le contrôle de la réponse immunitaire innée.

Dans une étude récente, deux types cellulaires de la barrière hémato-encéphalique, les astrocytes et les cellules endothéliales, ont été transduits avec des shARN contre les gènes impliqués dans l'AGS, et notamment SAMHD1 (Cuadrado et al., 2015). L'étude de ces cellules n'exprimant plus SAMHD1 a permis de montrer que la diminution de l'expression de SAMHD1 dans les cellules endothéliales est à l'origine d'une augmentation de l'expression de plusieurs ISG (IFIT2, IFIT3 et IRF9). Dans les astrocytes, la protéine SAMHD1 n'est pas détectable par immunoblot et la diminution de son expression n'a pas pu être évaluée. En revanche les shARN permettent bien une diminution de la quantité d'ARNm SAMHD1, ce qui induit la sécrétion des chémokines CXCL10 et CCL5 (également appelée RANTES).

Il a également été montré dès 2008 que l'expression de SAMHD1 dans les fibroblastes est augmentée après traitement par du TNF α et que l'abolition de cette induction de SAMHD1 limite l'induction de l'expression de différentes protéines d'adhésion et cytokines pro-inflammatoires (Liao et al., 2008).

Ces résultats illustrent l'importance de SAMHD1 dans la réponse immunitaire innée, à la fois comme protéine induite par certaines cytokines et comme protéine régulant l'induction de facteurs de l'immunité innée.

Cependant aucun mécanisme n'a été identifié pour expliquer la façon dont SAMHD1 intervient dans la réponse immunitaire innée. La principale hypothèse avancée est que l'expression de SAMHD1 permet de limiter l'accumulation d'acides nucléiques anormaux dans le cytoplasme et donc d'éviter l'activation de senseurs stimulant la réponse immunitaire innée. Cela expliquerait pourquoi la perte de SAMHD1 induit une hyper-inflammation et un phénotype comme l'AGS, en conduisant à l'accumulation de produits immunogènes. En revanche cela n'explique pas comment SAMHD1 favorise la production de certaines cytokines après traitement par le TNF α .

De plus, cette hypothèse reste très vague. En effet, aucune information n'est disponible sur le type d'acides nucléiques concernés (ADN ou ARN, simple ou double brin, possédant ou non certaines structures comme une coiffe ou une queue polyA...) ni sur l'activité de SAMHD1 impliquée dans ce phénotype. SAMHD1 dégrade-t-il ces acides nucléiques via son activité nucléase ou empêche-t-il leur synthèse via son activité dNTPase ?

Il serait sans doute plus prudent de considérer séparément l'effet de SAMHD1 sur la réponse immunitaire innée en général et son effet particulier dans la limitation de la mise en place de réponses immunitaires innées inadéquates, perdu chez les patients souffrant d'AGS.

Ainsi, de façon surprenante, si beaucoup de données dont certaines anciennes montrent que SAMHD1 joue un rôle clé dans la régulation de la réponse immunitaire innée, quasiment aucune information n'indique comment et dans quelles voies de signalisation cette protéine est impliquée.

Fonctions de SAMHD1 à l'échelle de l'organisme

Si relativement peu d'informations sont disponibles sur les mécanismes mis en jeu, trois fonctions cellulaires de SAMHD1 sont clairement identifiées à ce jour : un rôle dans la régulation du cycle cellulaire, dans la réponse à certains dommages à l'ADN et dans la régulation de la réponse immunitaire innée. L'implication de SAMHD1 dans une pathologie systémique comme l'AGS et la corrélation entre la perte de SAMHD1 et le développement de cancers soulignent que les fonctions cellulaires de cette protéine sont importantes à l'échelle de l'organisme.

SAMHD1 est exprimé dans la plupart des tissus humains, à l'exception notable du cerveau et du thymus (Li et al., 2000). Il semble cependant logique de considérer que les fonctions cellulaires de SAMHD1 n'ont pas la même importance dans tous les types cellulaires. Ainsi, son effet sur le cycle cellulaire n'est pas primordial dans les cellules ne se divisant pas alors que son effet sur la réponse aux dommages à l'ADN est important dans l'ensemble des cellules. De même, les fonctions cellulaires de SAMHD1 doivent logiquement varier dans le temps, son rôle dans la régulation de la production de cytokines étant plus important dans un contexte de réponse immunitaire qu'en l'absence de pathogène. La régulation de l'expression de SAMHD1 et de sa phosphorylation sur la T592 par l'IFN va dans le sens de cette hypothèse.

Ainsi, comprendre les mécanismes à l'origine des fonctions cellulaires de SAMHD1 serait une première étape vers la compréhension de ses fonctions à l'échelle de l'organisme. Cependant ces dernières ne correspondent vraisemblablement pas uniquement à l'articulation à différents niveaux des fonctions cellulaires de la protéine. En effet, au moins une fonction systémique non expliquée par les fonctions cellulaires de SAMHD1 a été mise en évidence à ce jour.

Comme expliqué en introduction du présent manuscrit, les patients atteints d'AGS dû à des mutations de SAMHD1 présentent un phénotype particulier de défauts dans la vascularisation cérébrale pouvant conduire à des hémorragies (Ramesh et al., 2010; Thiele et al., 2010; Xin et al., 2011). Si aucun effet de la perte de SAMHD1 n'a été observé sur l'angiogenèse in vitro (Cuadrado et al., 2015), les phénotypes de ces patients suggèrent tout de même un rôle de SAMHD1 dans l'angiogenèse cérébrale in vivo. De plus, même si l'expression de SAMHD1 est réduite dans les tissus cérébraux, l'observation de phénotypes encéphaliques chez les patients AGS et d'effets de shARN SAMHD1 dans les astrocytes (Cuadrado et al., 2015) montrent que SAMHD1 possède une fonction cérébrale chez l'homme.

Méconnaissance des mécanismes de régulation de SAMHD1

Si plusieurs articles récents, dont le mien, permettent de mieux comprendre les fonctions cellulaires de SAMHD1, le manque de données concernant la régulation de SAMHD1 constitue une grande limite au choix des modèles expérimentaux.

En termes de transcription, il est connu que le promoteur du gène SAMHD1 peut être inhibé par méthylation (De Silva et al., 2013a; De Silva et al., 2013b; Wang et al., 2014a), mais ce phénomène a pour l'instant essentiellement été mis en évidence dans le contexte de deux types de cancer et dans des lignées issues de leucémies. Dans l'étude réalisée sur des lignées cellulaires, une expérience a permis de montrer que le promoteur de SAMHD1 n'est pas méthylé dans les lymphocytes CD4+ primaires (De Silva et al., 2013b). Il s'agit du seul résultat publié concernant la méthylation du

promoteur de SAMHD1 in vivo en dehors d'un contexte pathologique. Il serait intéressant d'étudier le statut du promoteur SAMHD1 dans un panel plus large de types cellulaires.

En dehors de cette méthylation du promoteur, un seul résultat a été publié sur la présence de marques épigénétiques, activatrices ou inhibitrices, susceptible d'influencer la transcription de SAMHD1. Il a été montré que, dans certaines lignées cellulaires comme les Jurkat et les SupT1, le traitement avec un inhibiteur d'histone-déacétylases permet d'augmenter la transcription de SAMHD1 (De Silva et al., 2013b). Cela suggère que l'expression de SAMHD1 dans ces lignées est normalement inhibée par déacétylation des histones. Aucun résultat n'a été publié depuis concernant l'importance de cette régulation épigénétique dans l'expression de SAMHD1 in vivo.

Un certain nombre de résultats expérimentaux y compris les miens montrent qu'augmenter l'expression des ARNm SAMHD1 ne permet pas nécessairement d'augmenter la quantité de la protéine correspondante (De Silva et al., 2013b). Le seul mécanisme de régulation actuellement identifié entre l'étape de transcription et l'étape de traduction de SAMHD1 est la régulation par un miR (Jin et al., 2014). Mais aucune étude n'a été réalisée pour voir si ce miR suffit à expliquer l'ensemble des phénotypes observés ou si d'autres mécanismes de régulation sont mis en jeu. En effet, la mesure des ARNm SAMHD1 par qPCR montre que ceux-ci sont transcrits et stables, mais rien ne prouve que leur maturation et leur transport se font de façon normale. Or des perturbations dans ces étapes peuvent entraîner un défaut de traduction.

Les lignées Jurkat et SupT1 traitées avec des inhibiteurs de méthylation de l'ADN et de déacétylation des histones, dans lesquelles la transcription du gène SAMHD1 serait maximisée, pourraient constituer de bons modèles pour étudier les mécanismes de régulation intervenant entre la transcription et la traduction des ARNm SAMHD1.

Enfin, la régulation de SAMHD1 au niveau post-traductionnel est également mal comprise. Différents sites de phosphorylation ont été identifiés sur cette protéine et seul l'un d'entre eux, la T592, a réellement été étudié à l'heure actuelle. Mais même pour cette phosphorylation, beaucoup de questions restent à éclaircir, parmi lesquelles le rôle de CDK1 et de CDK2 dans la mise en place de cette phosphorylation, l'identification de phosphatase(s) susceptible(s) de la cliver et l'effet même de cette phosphorylation sur l'activité catalytique de SAMHD1. En effet, s'il est connu que la phosphorylation sur la T592 ne modifie pas l'activité dNTPase de SAMHD1, le mécanisme expliquant la perte de l'activité antivirale de cette protéine après phosphorylation n'est pas identifié.

Cette méconnaissance globale des mécanismes de régulation de SAMHD1, et notamment de ses activateurs et inhibiteurs au niveau post-traductionnel limite l'évaluation de l'importance de l'altération de SAMHD1 dans la tumorigenèse. En effet, nous n'avons observé une diminution de la quantité de la protéine SAMHD1 que chez une partie des patients cancéreux, mais il est possible que la fonction de SAMHD1 soit altérée chez un nombre plus important de patients, via l'inhibition d'un activateur ou l'activation d'un inhibiteur. Une meilleure compréhension du réseau de régulation de SAMHD1 permettrait d'évaluer avec plus de fiabilité la prévalence de la perte de SAMHD1 dans le développement de différents cancers. La perte de fonction de SAMHD1 étant susceptible de constituer un facteur de résistance à certaines chimiothérapies, il est important de pouvoir la détecter y compris lorsqu'elle est due à une perturbation de facteurs régulant SAMHD1.

Angles d'approches pour l'étude de SAMHD1

Les premiers articles concernant SAMHD1 datent du milieu des années 1990. Néanmoins, sur plus de 230 publications mentionnant cette protéine accessibles à ce jour, seules 20 sont antérieures à son identification en tant que facteur de restriction du VIH-1. Plus des trois quarts des articles ultérieurs concernent l'activité antivirale de cette protéine et/ou sa régulation par la protéine virale Vpx. Ces quelques chiffres illustrent le fait que la plupart des travaux effectués sur SAMHD1 se placent

dans le contexte d'une infection virale, par le VIH-1 ou un autre des virus restreints par SAMHD1. Parmi les publications restantes, un certain nombre concernent l'AGS. Ainsi, très peu de publications s'intéressent aux fonctions cellulaires de SAMHD1 et à sa régulation indépendamment de la protéine Vpx.

SAMHD1, un facteur de restriction

Les intérêts d'une meilleure compréhension de la façon dont SAMHD1 module l'infection de plusieurs types cellulaires ne se divisant pas ou peu par différents virus sont évidents.

D'une part, cela permet la mise au point de meilleurs outils expérimentaux et thérapeutiques. L'exemple le plus frappant est sans doute la nette amélioration de la transduction de certains types cellulaires comme les cellules dendritiques par des vecteurs lentiviraux et rétroviraux grâce à l'addition de la protéine virale Vpx. Ce progrès technique est d'ores et déjà utilisé en laboratoire, mais il pourrait également être utilisé dans des essais de thérapie génique.

D'autre part le VIH-1 est actuellement à l'origine d'une pandémie et tout élément améliorant la compréhension des mécanismes d'infection de ce pathogène est susceptible d'améliorer à terme les traitements disponibles pour les patients voire de contribuer à l'élaboration d'un vaccin efficace.

SAMHD1 est considéré comme pouvant constituer un élément clé de la pathogenèse du VIH-1 car l'incapacité de ce virus à contrecarrer SAMHD1 est une de ses principales différences avec le VIH-2. Or les patients infectés par le VIH-2 ont une réponse immunitaire plus efficace que celle des patients infectés par le VIH-1, aussi bien sur le plan cellulaire (Zheng et al., 2004a) que sur le plan humoral (Kong et al., 2012). Parallèlement, les patients infectés par le VIH-2 ont une charge virale significativement plus faible que celle observée chez les patients infectés par le VIH-1 dès lors que leur taux de lymphocytes CD4+ est supérieur à 200 millions par litre de sang (Simon et al., 1993). Or les patients infectés par le VIH-2 ont des taux de CD4+ plus élevés et une probabilité plus faible de développer un SIDA que les patients infectés par le VIH-1 (Marlink et al., 1994).

Sachant cela, identifier les mécanismes par lesquels SAMHD1 module la pathogenèse du VIH-1 paraît être la première étape pour établir des protocoles visant à les contrecarrer pour réduire cette pathogenèse et la rapprocher de celle du VIH-2.

De plus, SAMHD1 empêche l'infection des cellules dendritiques par le VIH-1 et il a été montré que rendre ces cellules sensibles au VIH-1 induit leur maturation et promeut une meilleure réponse immunitaire antivirale via la production d'IFN de type I et l'activation des lymphocytes T (Manel et al., 2010). Dans cette optique, administrer la protéine virale Vpx dans des essais vaccinaux contre le VIH-1 devrait permettre d'augmenter leur efficacité et peut-être d'arriver à concevoir un vaccin efficace contre le VIH-1.

Si cela ne remet pas en cause la possibilité d'améliorer les vaccins anti-VIH-1 grâce à la protéine virale Vpx, des données expérimentales suggèrent que la capacité des patients infectés par le VIH-2 à contrôler leur infection n'est pas liée à la dégradation de SAMHD1 par ce virus. En effet, il a été montré que la capacité des protéines Vpx à contrecarrer SAMHD1 est identique qu'elles soient isolées chez des patients contrôlant l'infection par le VIH-2 ou chez des patients ayant développé un SIDA après infection par le VIH-2 (Yu et al., 2013). Ainsi la capacité à dégrader SAMHD1 n'est pas un déterminant majeur de l'évolution de la pathogenèse du VIH-2. A l'inverse, les seuls Vpx incapables de dégrader SAMHD1 ont été isolés chez des patients contrôlant l'infection par le VIH-2 même si rien ne permet d'exclure quand dans ces souches virales cette fonction soit remplie par la protéine virale Vpr.

Ces dernières données montrent clairement que comprendre les mécanismes par lesquels le VIH-2 contrecarre SAMHD1 et les reproduire dans le cas d'une infection par le VIH-1 a peu de chance de constituer une solution miracle pour éviter l'évolution vers le SIDA.

SAMHD1, un garant du fonctionnement cellulaire

L'identification de SAMHD1 en tant que facteur de restriction du VIH-1 a permis d'intéresser de nombreuses équipes de recherche à cette protéine. Néanmoins je pense que cette identification originelle est aujourd'hui un obstacle à l'étude de SAMHD1. En effet, ce caractère antiviral n'est que l'arbre qui cache la forêt. A l'heure actuelle, je pense qu'il serait plus pertinent d'investir du temps, de la force de travail et de l'argent pour comprendre les grandes lignes des fonctions et de la régulation cellulaires de SAMHD1 que pour comprendre le rôle d'acides aminés précis dans l'interaction avec la protéine virale Vpx. Et cela que la pertinence soit définie en termes d'apports à la biologie de façon générale ou en termes de bénéfices pour des patients.

En effet, comme exposé précédemment, SAMHD1 intervient dans plusieurs processus cellulaires clés dont la dérégulation peut favoriser l'apparition de cancers. De plus, cette protéine joue un rôle essentiel dans la régulation de la réponse immunitaire innée chez l'homme puisque des mutations innées dans le gène correspondant causent un syndrome hyper-inflammatoire entraînant le décès précoce des nourrissons.

Or, en dépit de la mise en valeur de cette protéine par son identification en tant que facteur de restriction il y a environ quatre ans, très peu de choses sont connues sur son fonctionnement et sa régulation indépendamment de Vpx. Seule une de ses activités catalytiques est clairement caractérisée à ce jour et peu de choses sont connues sur sa régulation. Il y a encore moins de résultats disponibles concernant ses mécanismes d'action dans la cellule.

Au-delà des lacunes déjà mentionnées dans ce manuscrit, certaines données publiées dans les études antérieures à 2011 n'ont pas été plus finement étudiées depuis alors qu'elles sont importantes pour comprendre la régulation de SAMHD1.

Par exemple, il a été publié en 2008 que le traitement de fibroblastes par du $\text{TNF}\alpha$ induit l'expression de SAMHD1 et que cette induction est médiée par le facteur de transcription IRF-1 (Liao et al., 2008). Or cette protéine dont l'expression est induite par l'IFN est connue depuis un certain temps comme étant un suppresseur de tumeur limitant la prolifération cellulaire et pouvant induire l'apoptose, notamment en cas de dommages importants à l'ADN (Kröger et al., 2002). De plus, IRF-1 est impliqué dans différentes étapes de la réponse immunitaire (Kröger et al., 2002). Ces fonctions connues d'IRF-1 ne sont pas sans rappeler celles de SAMHD1, et le lien entre ces deux protéines a été identifié en 2008. Pourtant aucune étude à ce jour n'a été réalisée sur le rôle d'IRF-1 dans la régulation de SAMHD1 ou sur l'importance éventuelle de SAMHD1 dans les fonctions cellulaires d'IRF-1.

Enfin, certains résultats récents suggèrent que SAMHD1 possède des fonctions cellulaires qui n'ont pas encore été caractérisées à ce jour. Ainsi il a été montré que SAMHD1 interagit avec eEF1A1 (Morrissey et al., 2015). Cette protéine est un facteur d'initiation de la traduction, mais elle joue également un rôle dans le contrôle qualité des protéines et leur dégradation. Des co-immunoprécipitations ont permis de montrer que SAMHD1 et eEF1A1 interagissent avec Rbx1 et la culline 4A dans des cellules THP-1 (Morrissey et al., 2015). Compte-tenu de ces résultats, il serait intéressant d'étudier les effets d'une déplétion ou d'une sur-expression de SAMHD1 sur la traduction et sur le contrôle qualité des protéines cellulaires.

Conclusion générale

Dans la recherche scientifique, il arrive qu'une découverte réalisée dans le cadre d'un domaine disciplinaire vienne enrichir un autre domaine disciplinaire. Ainsi, l'étude des interactions entre les pathogènes et leurs hôtes a souvent été à l'origine de découvertes majeures concernant la biologie de ces hôtes. Celles-ci pouvant concerner aussi bien des phénomènes systémiques que des acteurs moléculaires.

Mon doctorat s'inscrit dans cette démarche. En effet, peu de temps avant le début de celui-ci, la protéine SAMHD1 a été identifiée comme un facteur de restriction du VIH-1 dans les cellules de la lignée myéloïde. L'objectif de ma thèse était de mieux comprendre les fonctions cellulaires et les mécanismes de régulation de cette protéine, sur laquelle très peu de choses étaient connues à l'époque. Cet objectif s'est articulé autour de quatre projets indépendants.

L'un de ces projets concernait l'étude de la régulation de la protéine Vpx du VIH-2, elle-même étant un régulateur de SAMHD1. Deux autres portaient sur l'étude de deux aspects différents de la régulation de SAMHD1, potentiellement indépendants de l'infection par le VIH. Enfin, le projet ayant abouti à la publication d'un article et qui fait l'objet du présent manuscrit concernait le déchiffrement des fonctions cellulaires de SAMHD1.

J'ai pu montrer dans le cadre de ce dernier projet que SAMHD1 est impliqué dans des processus cellulaires aussi fondamentaux que la prolifération et la réparation des ruptures double brin de l'ADN. J'ai également observé une forte corrélation entre la perte de cette protéine et le développement de différents types de cancers. Ces résultats ont pour certains été confirmés par d'autres équipes et de plus en plus de travaux insistent sur le rôle de SAMHD1 dans la régulation de la réponse immunitaire.

L'ensemble de ces données identifie SAMHD1 non pas simplement comme un facteur de restriction mais comme un acteur clé de l'homéostasie cellulaire dont la perturbation est susceptible de causer des maladies graves comme le cancer ou l'AGS.

Ainsi, dans le cadre de mon doctorat, j'ai contribué à souligner l'importance fonctionnelle de SAMHD1 en dehors du contexte d'infection par le VIH. J'espère qu'à leur échelle ces travaux contribueront à attirer l'intérêt d'équipes de recherche sur SAMHD1 au-delà du domaine de la virologie.

Annexes

Descriptif des autres projets menés au cours du doctorat

Les quatre projets sur lesquels j'ai travaillé au cours de mon doctorat s'intègrent dans deux thématiques distinctes concernant la protéine SAMHD1. En effet, le projet présenté dans ce manuscrit concerne l'étude des fonctions cellulaires de SAMHD1 alors que les trois autres sont en lien avec la régulation de SAMHD1 par la protéine virale Vpx (Figure 35).

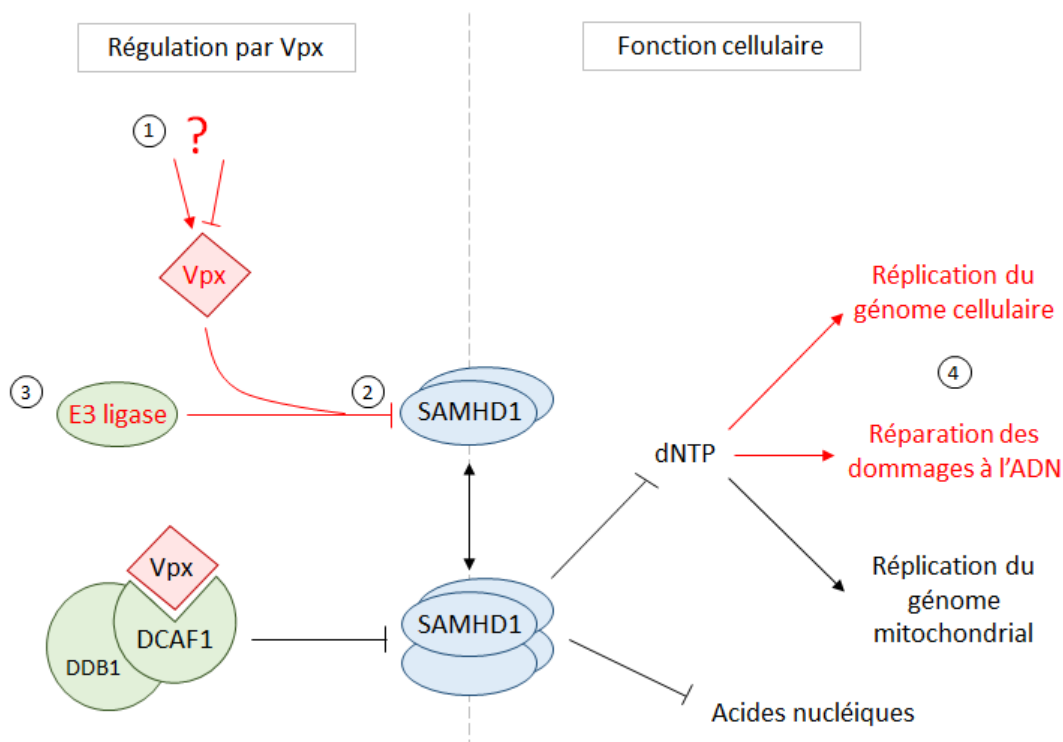


Figure 35 : schéma simplifié de l'état des connaissances concernant la régulation de SAMHD1 par Vpx et les fonctions cellulaires de SAMHD1. Les points abordés au cours de mon travail de thèse sont indiqués en rouge. Leur numérotation correspond à celle utilisée dans cette partie du manuscrit.

1/ Etude de la régulation de la protéine Vpx

La régulation de SAMHD1 par Vpx est caractérisée de façon de plus en plus précise, mais peu de données sont disponibles concernant les mécanismes de régulation de la protéine Vpx elle-même. Des expériences préliminaires nous ont permis de constater que la protéine Vpx du virus VISmac251 a une demi-vie très courte et peut être stabilisée par traitement avec un inhibiteur du protéasome. Nous avons alors souhaité identifier l'ubiquitine ligase impliquée dans la dégradation de la protéine virale Vpx.

Pour cela nous avons utilisé la liste des partenaires de Vpx préalablement identifiés au laboratoire par spectrométrie de masse (Laguette et al., 2011). En effet, cette liste contient plusieurs ubiquitine ligases. L'expression de chacune de ces ubiquitine ligases a été supprimée sélectivement par ARN interférence dans des cellules 293T et l'effet de cette suppression sur la stabilité de la protéine Vpx a été évalué. Nous avons ainsi identifié deux ubiquitine ligases dont la dégradation stabilise Vpx.

L'effet de la dégradation d'une de ces deux ubiquitine ligases a été mesuré sur différentes souches de Vpx disponibles au laboratoire. Nous avons ainsi dans un premier temps observé que les protéines Vpx de virus simiens avaient une demi-vie très inférieure à celle des protéines Vpx de virus infectant l'Homme. De plus, la déplétion de l'ubiquitine ligase humaine stabilise les Vpx d'origine simienne mais pas les Vpx d'origine humaine. Des expériences préliminaires réalisées avec la seconde

ubiquitine ligase dont la dégradation stabilise Vpx semblent mettre en évidence les mêmes phénotypes.

Ces données suggèrent que les Vpx des virus adaptés à l'Homme ont évolué pour échapper aux mécanismes de dégradation actifs contre les Vpx de virus simiens dans les cellules humaines. Cela aurait abouti à une augmentation la demi-vie de la protéine Vpx des virus infectant l'Homme dans les cellules humaines. Ce projet n'est cependant pas terminé. En effet, les rôles respectifs des deux ubiquitine ligases identifiées dans la dégradation des Vpx d'origine simienne restent à quantifier. De plus, l'ubiquitination directe de Vpx par ces ubiquitine ligases n'a pas encore été mise en évidence. Enfin, il serait intéressant de mesurer la stabilité des Vpx de virus simiens et humains dans des cellules simiennes et d'étudier l'effet des homologues simiens des ubiquitine ligases préalablement identifiées chez l'Homme sur leur demi-vie.

2/ Etude de l'effet de Vpx sur l'expression à long terme de SAMHD1

En mesurant la demi-vie de certaines souches de Vpx, nous avons observé que le Vpx du VISmac251 a une demi-vie de l'ordre d'un quart d'heure. Or le traitement de cellules THP-1 avec des particules pseudo-virales contenant cette protéine induit une dégradation de la protéine SAMHD1 qui persiste pendant plusieurs jours. Autrement dit, le traitement de cellules THP-1 avec la protéine Vpx induit une élimination de la protéine SAMHD1 qui persiste même après dégradation de la protéine Vpx. Nous avons dans un premier temps essayé de caractériser plus finement ce phénotype et dans un second temps testé différentes hypothèses pour identifier le mécanisme mis en jeu.

Ce projet étant encore en cours au laboratoire, les détails n'en seront pas présentés dans ce manuscrit.

3/ Recherche d'autres ubiquitine ligases recrutées par Vpx

Avant même l'identification de SAMHD1 comme cible de ce mécanisme, Vpx était connu comme étant capable de recruter la CRL4 pour ubiquitiner un facteur cellulaire et induire sa dégradation par le protéasome. Néanmoins, dans certains types cellulaires comme les cellules dendritiques et les cellules THP-1 différenciées, la capacité des souches de Vpx à interagir avec les composants de la CRL4 ne corrèle pas avec leur capacité à augmenter l'infection par le VIH-1 (Goujon et al., 2007). Cette observation suggère que, au moins dans ces types cellulaires, une autre ubiquitine ligase est recrutée par Vpx pour induire la dégradation de SAMHD1.

Afin d'identifier cette ubiquitine ligase, nous sommes à nouveau partis de la liste des ubiquitine ligases préalablement identifiées comme interagissant avec Vpx. L'expression de chacune de ces ubiquitine ligases a été diminuée par ARN interférence dans des cellules THP-1 différenciées. Les cellules correspondantes ont ensuite été traitées avec des particules pseudo-virales contenant la protéine Vpx ou des particules contrôles puis infectées avec un VIH-1 exprimant le gène rapporteur GFP. L'effet de la diminution de la quantité des différentes ubiquitine ligases sur l'augmentation de l'infection par la protéine Vpx a ainsi été évalué. Nous avons pu observer une diminution voire une abolition de l'effet de Vpx sur l'infection par le VIH-1 suite à l'extinction de plusieurs de ces ubiquitine ligases.

Pour s'assurer que ce phénotype était bien lié à l'effet de Vpx sur la dégradation de SAMHD1, la même expérience a été reproduite dans des cellules HeLa, dans lesquelles SAMHD1 n'est pas actif contre le VIH-1. Parmi les ubiquitine ligases préalablement identifiées, une seule n'avait aucun effet sur l'infection des cellules HeLa par le VIH-1, en l'absence et en présence de Vpx.

En résumé, cette ubiquitine ligase possède les caractéristiques suivantes : i/ sa suppression dans les cellules THP-1 différenciées, où SAMHD1 est actif, abolit l'effet de Vpx sur l'infection par le VIH-1 ii/ sa suppression dans les cellules HeLa, où SAMHD1 est inactif, n'a aucun effet sur l'infection par le VIH-1 iii/ elle interagit avec Vpx. Ces propriétés en font un candidat potentiellement recruté par Vpx pour induire la dégradation de SAMHD1 dans les THP-1.

Après des premiers résultats encourageants, ce projet a finalement été mis de côté pour des raisons techniques. Sa progression aurait nécessité de longues mises au point qui n'étaient pas compatibles avec la durée restante de doctorat et la priorité a donc été donnée au projet numéro 2.

Matériel et méthodes supplémentaire

Cette section comprend les informations techniques relatives aux figures présentées comme « résultats additionnels » dans le présent manuscrit.

Traitements des cellules

Les différentes lignées cellulaires ont été traitées avec du Mg132 (Sigma, M7449) à une concentration de 10 μ M pendant quatre heures. Les cellules HeLa surexprimant stablement une construction SAMHD1 sauvage ont été traitées avec du méthanesulfonate de méthyle (Sigma, 129925) à une concentration de 10 μ g/mL pendant quatre heures.

Séquençage

Les produits PCR d'intérêt ont été purifiés à partir d'un gel d'agarose 1% à l'aide du kit QIAex extraction (QIAgen) en suivant les instructions du fabricant. Ils ont ensuite été insérés dans le vecteur pCR-BluntII-TOPO (Life technologies) conformément au protocole du fabricant. Après ligation, ces plasmides ont été transformés dans des bactéries TOP 10 (Invitrogen) par un choc thermique de 30 secondes à 42°C. Ces bactéries ont été séparées sur des boîtes de gélose nutritive LB. Des colonies ont été obtenues après 15 heures de croissance à 37°C et cinq d'entre elles pour chaque lignée cellulaire ont été récupérées. Leur ADN plasmidique a été purifié en utilisant le kit QIAprep Spin Miniprep (QIAgen) conformément aux instructions du fabricant puis séquencé.

Méthodes précédemment décrites

Les extraits cellulaires totaux et nucléaires, les immunoblots et les gradients glycérol ont été réalisés avec les méthodes présentées dans l'article. Il en est de même pour les PCR quantitatives et l'analyse statistique des résultats obtenus.

Echantillons

Les ARNs de 40 patients atteints de cancers hématopoïétiques ont été obtenus auprès de la tumorotheque de l'Hôpital Saint-Louis.

Références bibliographiques

- Abercrombie, M. (1967). Contact inhibition: the phenomenon and its biological implications. *Natl. Cancer Inst. Monogr.* 26, 249–277.
- Adrain, C. and Martin, S.J. (2001). The mitochondrial apoptosome: A killer unleashed by the cytochrome seas. *Trends Biochem. Sci.* 26, 390–397.
- Agarwal, N., Iyer, D., Patel, S., Sekhar, R., Philips, T., Schubert, U., Buras, T.O.E., Samson, S., Couturier, J., Lewis, D., Rodriguez-Barradas, M., Jahoor, F., Kino, T., Kopp, J., and Balasubramanyam, A. (2013). HIV-1 Vpr Induces Adipose Dysfunction in Vivo Through Reciprocal Effects on PPAR/GR Co-Regulation. *Sci Transl Med.*
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2008). *Molecular Biology of the Cell*, Fifth Edition.
- Allouch, A., David, A., Amie, S.M., Lahouassa, H., Chartier, L., Margottin-Goguet, F., Barré-Sinoussi, F., Kim, B., Sáez-Cirión, A., and Pancino, G. (2013). p21-mediated RNR2 repression restricts HIV-1 replication in macrophages by inhibiting dNTP biosynthesis pathway. *Proc. Natl. Acad. Sci. U. S. A.* 110, E3997–E4006.
- Ambrosio, S., Amente, S., Napolitano, G., Di Palo, G., Lania, L., and Majello, B. (2015). MYC impairs resolution of site-specific DNA double-strand breaks repair. *Mutat. Res. Mol. Mech. Mutagen.* 774, 6–13.
- Amie, S.M., Bambara, R. a., and Kim, B. (2013). GTP is the primary activator of the anti-HIV restriction factor SAMHD1. *J. Biol. Chem.* 288, 25001–25006.
- Aravind, L. and Koonin, E. V. (1998). The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem. Sci.* 23, 469–472.
- Arhel, N. (2010). Revisiting HIV-1 uncoating. *Retrovirology* 7, 96.
- Avery, O.T., MacLeod, C.M., and McCarty, M. (1945). Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types. *J. Nerv. Ment. Dis.* 102, 213–214.
- Baldauf, H.-M., Pan, X., Erikson, E., Schmidt, S., Daddacha, W., Burggraf, M., Schenkova, K., Ambiel, I., Wabnitz, G., Gramberg, T., Panitz, S., Flory, E., Landau, N.R., Sertel, S., Rutsch, F., Lasitschka, F., Kim, B., König, R., Fackler, O.T., and Keppler, O.T. (2012). SAMHD1 restricts HIV-1 infection in resting CD4+ T cells. *Nat. Med.* 18, 1682–1689.
- Balducci, L. and Ershler, W.B. (2005). Cancer and ageing: a nexus at several levels. *Nat. Rev. Cancer* 5, 655–662.
- Baltimore, D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* 226, 1209–1211.
- Baltimore, D. and Franklin, R.M. (1962). Preliminary Data on a Virus-Specific Enzyme Responsible for Synthesis of Viral RNA. *Biochem. Biophys. Res. Commun.* 9.

- Baltimore, D., Franklin, R.M., Tamm, I., and Eggers, H.J. (1963). Studies on Enzymatic Mechanism of Virus RNA Synthesis. *Fed. Proc.* 22.
- Barré-sinoussi, F. (2003). The early years of HIV research : integrating clinical and basic research. *Nat. Med.* 9, 844–846.
- Barre-Sinoussi, F., J. C. Chermann, E. Al (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* (80-). 220(4599), 868–871.
- Batista, L.F.Z. (2014). *Telomere biology in stem cells and reprogramming* (Elsevier Inc.).
- Beer, B.E., Bailes, E., Sharp, P.M., and Hirsch, V.M. (1999). Diversity and Evolution of Primate Lentiviruses. *Hum. Retroviruses AIDS 1999 a Compil. Anal. Nucleic Acid Amin. Acid Seq.* 460–474.
- Behrendt, R., Schumann, T., Gerbaulet, A., Nguyen, L. a., Schubert, N., Alexopoulou, D., Berka, U., Lienenklaus, S., Peschke, K., Gibbert, K., Wittmann, S., Lindemann, D., Weiss, S., Dahl, A., Naumann, R., Dittmer, U., Kim, B., Mueller, W., Gramberg, T., and Roers, A. (2013). Mouse SAMHD1 has antiretroviral activity and suppresses a spontaneous cell-intrinsic antiviral response. *Cell Rep.* 4, 689–696.
- Beignon, A., Mckenna, K., Skoberne, M., Manches, O., Dasilva, I., Kavanagh, D.G., Larsson, M., Gorelick, R.J., Lifson, J.D., and Bhardwaj, N. (2005). Endocytosis of HIV-1 activates plasmacytoid dendritic cells via Toll-like receptor – viral RNA interactions. 115.
- Beloglazova, N., Flick, R., Tchigvintsev, A., Brown, G., Popovic, A., Nocek, B., and Yakunin, A.F. (2013). Nuclease activity of the human SAMHD1 protein implicated in the aicardi-gouti???res syndrome and HIV-1 restriction. *J. Biol. Chem.* 288, 8101–8110.
- Berg, R.K., Melchjorsen, J., Rintahaka, J., Diget, E., S by, S., Kristy, A., Gorelick, R.J., Matikainen, S., Larsen, C.S., Ostergaard, L., S ren, R., and Mogensen, T.H. (2012). Genomic HIV RNA Induces Innate Immune Responses through RIG-I-Dependent Sensing of Secondary- Structured RNA. 7, 1–10.
- Berger, A., Sommer, A.F.R., Zwarg, J., Hamdorf, M., Welzel, K., Esly, N., Panitz, S., Reuter, A., Ramos, I., Jatiani, A., Mulder, L.C.F., Fernandez-Sesma, A., Rutsch, F., Simon, V., K nig, R., and Flory, E. (2011a). SAMHD1-deficient CD14+ cells from individuals with Aicardi-Gouti res syndrome are highly susceptible to HIV-1 infection. *PLoS Pathog.* 7, 1–12.
- Berger, G., Goujon, C., Darlix, J.-L., and Cimorelli, a (2009). SIVMAC Vpx improves the transduction of dendritic cells with nonintegrative HIV-1-derived vectors. *Gene Ther.* 16, 159–163.
- Berger, G., Durand, S., Fargier, G., Nguyen, X.N., Cordeil, S., Bouaziz, S., Muriaux, D., Darlix, J.L., and Cimorelli, A. (2011b). Apobec3a is a specific inhibitor of the early phases of hiv-1 infection in myeloid cells. *PLoS Pathog.* 7.
- Bhome, R., Bullock, M.D., Al Saihati, H. a., Goh, R.W., Primrose, J.N., Sayan, a. E., and Mirnezami, A.H. (2015). A top-down view of the tumor microenvironment: structure, cells and signaling. *Front. Cell Dev. Biol.* 3, 1–9.
- Bieda, K., Hoffmann, a, and Boller, K. (2001). Phenotypic heterogeneity of human endogenous retrovirus particles produced by teratocarcinoma cell lines. *J. Gen. Virol.* 82, 591–596.

- Bishop, K.N., Verma, M., Kim, E.Y., Wolinsky, S.M., and Malim, M.H. (2008). APOBEC3G inhibits elongation of HIV-1 reverse transcripts. *PLoS Pathog.* 4, 13–20.
- Blaise, S., de Parseval, N., Bénit, L., and Heidmann, T. (2003). Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. *Proc. Natl. Acad. Sci. U. S. A.* 100, 13013–13018.
- Boye, E., Skjölberg, H., and Grallert, B. (2009). Checkpoint Regulation of DNA Replication. In *DNA Replication, Methods and Protocols*, S. Vengrova, and J. Daalgard, eds. (Springer Protocols), pp. 493–507.
- Bray, M., Prasad, S., Dubay, J.W., Hunter, E., Jeang, K.T., Rekosh, D., and Hammarskjöld, M.L. (1994). A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. *Proc. Natl. Acad. Sci. U. S. A.* 91, 1256–1260.
- Brennan, C.M., Gallouzi, I.E., and Steitz, J. a. (2000). Protein ligands to HuR modulate its interaction with target mRNAs in vivo. *J. Cell Biol.* 151, 1–13.
- Burma, S., Chen, B.P.C., and Chen, D.J. (2006). Role of non-homologous end joining (NHEJ) in maintaining genomic integrity. *DNA Repair (Amst.)* 5, 1042–1048.
- Burns, J.C., Friedmann, T., Driever, W., Burrascano, M., and Yee, J.K. (1993). Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 90, 8033–8037.
- Butel, J.S. and Fan, H. (2012). The diversity of human cancer viruses. *Curr. Opin. Virol.* 2, 449–452.
- Cao, W., Bover, L., Cho, M., Wen, X., Hanabuchi, S., Bao, M., Rosen, D.B., Wang, Y.-H., Shaw, J.L., Du, Q., Li, C., Arai, N., Yao, Z., Lanier, L.L., and Liu, Y.-J. (2009). Regulation of TLR7/9 responses in plasmacytoid dendritic cells by BST2 and ILT7 receptor interaction. *J. Exp. Med.* 206, 1603–1614.
- Capasso, C., Garofalo, M., Hirvinen, M., and Cerullo, V. (2014). The evolution of adenoviral vectors through genetic and chemical surface modifications. *Viruses* 6, 832–855.
- Carr, A.M. and Lambert, S. (2013). Replication stress-induced genome instability: The dark side of replication maintenance by homologous recombination. *J. Mol. Biol.* 425, 4733–4744.
- Casartelli, N., Sourisseau, M., Feldmann, J., Guivel-Benhassine, F., Mallet, A., Marcelin, A.G., Guatelli, J., and Schwartz, O. (2010). Tetherin restricts productive HIV-1 cell-to-cell transmission. *PLoS Pathog.* 6.
- Chahroudi, a., Bosinger, S.E., Vanderford, T.H., Paiardini, M., and Silvestri, G. (2012). Natural SIV Hosts: Showing AIDS the Door. *Science* (80-.). 335, 1188–1193.
- Chattergoon, M.A., Latanich, R., Quinn, J., Winter, M.E., Buckheit, R.W., Blankson, J.N., Pardoll, D., and Cox, A.L. (2014). HIV and HCV Activate the Inflammasome in Monocytes and Macrophages via Endosomal Toll-Like Receptors without Induction of Type 1 Interferon. *10*, 1–12.
- Chen, Z., Zhu, M., Pan, X., Zhu, Y., Yan, H., Jiang, T., Shen, Y., Dong, X., Zheng, N., Lu, J., Ying, S., and Shen, Y. (2014). Inhibition of Hepatitis B virus replication by SAMHD1. *Biochem. Biophys. Res. Commun.* 450, 1462–1468.

- Chinnam, M. and Goodrich, D.W. (2011). RB1, development and cancer. *Curr Tol Dev Biol* 94, 129–169.
- Chipuk, J.E. and Green, D.R. (2006). Dissecting p53-dependent apoptosis. *Cell Death Differ.* 13, 994–1002.
- Ciron, C., Desmaris, N., Colle, M.A., Raoul, S., Joussemet, B., Vérot, L., Ausseil, J., Froissart, R., Roux, F., Chérel, Y., Ferry, N., Lajat, Y., Schwartz, B., Vanier, M.T., Maire, I., Tardieu, M., Mouiller, P., and Heard, J.M. (2006). Gene therapy of the brain in the dog model of Hurler’s syndrome. *Ann. Neurol.* 60, 204–213.
- Clavel, F., Guyader, M., Guétard, D., Sallé, M., Montagnier, L., and Alizon, M. (1986). Molecular cloning and polymorphism of the human immune deficiency virus type 2. *Nature.* 324, 691–695.
- Clouse, K.N., Luo, M.J., Zhou, Z., and Reed, R. (2001). A Ran-independent pathway for export of spliced mRNA. *Nat. Cell Biol.* 3, 97–99.
- Cocka, L.J. and Bates, P. (2012). Identification of Alternatively Translated Tetherin Isoforms with Differing Antiviral and Signaling Activities. *PLoS Pathog.* 8.
- Cohen, M. et al. (2012). Prevention of HIV-1 Infection with Early Antiretroviral Therapy Myron. *N Engl J Med.* 365, 493–505.
- Cohen, C.J., Lock, W.M., and Mager, D.L. (2009). Endogenous retroviral LTRs as promoters for human genes: A critical assessment. *Gene* 448, 105–114.
- Consortium, I.H.G. (2002). Initial sequencing and analysis of the human genome. *Nature* 420, 520–562.
- Cordaux, R. and Batzer, M. (2009). The impact of retrotransposons on human genome evolution. *Nat Rev Genet.* 10, 691–703.
- Cornelis, G., Heidmann, O., Degrelle, S. a, Vernochet, C., Lavialle, C., Letzelter, C., Bernard-Stoecklin, S., Hassanin, A., Mulot, B., Guillomot, M., Hue, I., Heidmann, T., and Dupressoir, A. (2013). Captured retroviral envelope syncytin gene associated with the unique placental structure of higher ruminants. *Proc. Natl. Acad. Sci. U. S. A.* 110, E828–E837.
- Cornelis, G., Heidmann, O., Bernard-Stoecklin, S., Reynaud, K., Veron, G., Mulot, B., Dupressoir, a., and Heidmann, T. (2012). PNAS Plus: From the Cover: Ancestral capture of syncytin-Car1, a fusogenic endogenous retroviral envelope gene involved in placentation and conserved in Carnivora. *Proc. Natl. Acad. Sci.* 109, E432–E441.
- Cornelis, G., Vernochet, C., Malicorne, S., Souquere, S., Tzika, a. C., Goodman, S.M., Catzeflis, F., Robinson, T.J., Milinkovitch, M.C., Pierron, G., Heidmann, O., Dupressoir, a., and Heidmann, T. (2014). Retroviral envelope syncytin capture in an ancestrally diverged mammalian clade for placentation in the primitive Afrotherian tenrecs. *Proc. Natl. Acad. Sci.* 111, E4332–E4341.
- Cornelis, G., Vernochet, C., Carradec, Q., Souquere, S., Mulot, B., Catzeflis, F., Nilsson, M. a., Menzies, B.R., Renfree, M.B., Pierron, G., Zeller, U., Heidmann, O., Dupressoir, A., and Heidmann, T. (2015). Retroviral envelope gene captures and syncytin exaptation for placentation in marsupials. *Proc. Natl. Acad. Sci.* 112, E487–E496.

Cory, S. and Adams, J.M. (2002). The Bcl2 family: regulators of the cellular life-or-death switch. *Nat. Rev. Cancer* 2, 647–656.

Costiniuk, C.T. and Jenabian, M. -a. (2014). Cell-to-cell transfer of HIV infection: implications for HIV viral persistence. *J. Gen. Virol.* 95, 2346–2355.

Cowell, J.K., Hungerford, J., Jay, M., and Rutland, P. (1988). Retinoblastoma- clinical and genetic aspects: a review. *J. R. Soc. Med.* 81, 220–223.

Craigie, R. (2001). HIV Integrase, a Brief Overview from Chemistry to Therapeutics. *J. Biol. Chem.* 276, 23213–23216.

Cribier, A., Descours, B., Valadão, A., Laguette, N., and Benkirane, M. (2013). Phosphorylation of SAMHD1 by Cyclin A2/CDK1 Regulates Its Restriction Activity toward HIV-1. *Cell Rep.* 3, 1036–1043.

Crick, F. (1958). On protein synthesis. *Symp. Soc. Exp. Biol.* 12, 138–163.

Crick, F.H.C. (1956). On protein synthesis. *Symp. Soc. Exp. Biol. XII*, 139–163.

Crick, F.H.C. (1970). Central Dogma of Molecular Biology. *Nature* 227, 561–563.

Crow, Y.J., Chase, D.S., Schmidt, J.L., Szykiewicz, M., Forte, G.M.A., Gornall, H.L., Oojageer, A., Anderson, B., Pizzino, A., Helman, G., Abdel-hamid, M.S., Abdel-salam, G.M., Ackroyd, S., Aeby, A., Agosta, G., Albin, C., Allon-shalev, S., Arellano, M., Ariaudo, G., Aswani, V., Babul-hirji, R., Baildam, E.M., Bahi-buisson, N., Bailey, K.M., Barnerias, C., Barth, M., Battini, R., Bianchi, M., De, T.B., Beresford, M.W., Blair, E.M., Bloom, M., Burlina, A.B., Carpanelli, M.L., Carvalho, D.R., Castro-gago, M., Cavallini, A., Cereda, C., Chandler, K.E., Chitayat, D.A., Collins, A.E., Corcoles, C.S., Cordeiro, N.J. V., Crichtiutti, G., Dabydeen, L., Dale, R.C., Arrigo, S.D., Goede, C.G.E.L. De, Laet, C. De, Waele, L.M.H. De, Denzler, I., Desguerre, I., Devriendt, K., Rocco, M. Di, Fahey, M.C., Gener, B., Goizet, C., Gowrinathan, N.R., Gowrishankar, K., Hanrahan, D., Isidor, B., Khan, N., King, M.D., Kirk, E.P., Kumar, R., Lagae, L., Lin, J.S., Linnankivi, T., Mackay, M.T., Marom, D.R., Lourenc, C.M., Mckee, S.A., Moroni, I., Morton, J.E. V, Suri, M., Tacke, U., Tan, T.Y., Naude, W., Teik, K.W., Thomas, M.M., Till, M., Tonduti, D., Valente, E.M., Coster, R.N. Van, Knaap, M.S. Van Der, Vassallo, G., Vijzelaar, R., Vogt, J., Wallace, G.B., Wassmer, E., Webb, H.J., Whitehouse, W.P., Whitney, R.N., Zaki, M.S., Zuberi, S.M., Livingston, J.H., Rozenberg, F., Lebon, P., Vanderver, A., Orcesi, S., and Rice, G.I. (2015). Characterization of Human Disease Phenotypes Associated with Mutations in TREX1, RNASE H2A , RNASEH2B, RNASEH2C, SAMHD1, ADAR and IFIH1. *Am. J. Med. Genet.* 296–312.

Cuadrado, E., Michailidou, I., van Bodegraven, E.J., Jansen, M.H., Sluijs, J. a., Geerts, D., Couraud, P.-O., De Filippis, L., Vescovi, a. L., Kuijpers, T.W., and Hol, E.M. (2015). Phenotypic Variation in Aicardi-Goutieres Syndrome Explained by Cell-Specific IFN-Stimulated Gene Response and Cytokine Release. *J. Immunol.*

D’arc, M., Ayoub, A., Esteban, A., Learn, G.H., Boué, V., Liegeois, F., Etienne, L., Tagg, N., Leendertz, F.H., Boesch, C., Madinda, N.F., Robbins, M.M., Gray, M., Cournil, A., Ooms, M., Letko, M., Simon, V. a., Sharp, P.M., Hahn, B.H., Delaporte, E., Mpoudi Ngole, E., and Peeters, M. (2015). Origin of the HIV-1 group O epidemic in western lowland gorillas. *Proc. Natl. Acad. Sci.* 201502022.

Dale, R.C., Gornall, H., Singh-Grewal, D., Alcausin, M., Rice, G.I., and Crow, Y.J. (2010). Familial Aicardi-Goutières syndrome due to SAMHD1 mutations is associated with chronic arthropathy and contractures. *Am. J. Med. Genet. Part A* 152, 938–942.

- DeFalco, J., Tomishima, M., Liu, H., Zhao, C., Cai, X., Marth, J.D., Enquist, L., and Friedman, J.M. (2001). Virus-assisted mapping of neural inputs to a feeding center in the hypothalamus. *Science* 291, 2608–2613.
- Delelis, O., Carayon, K., Saïb, A., Deprez, E., and Mouscadet, J.-F. (2008). Integrase and integration: biochemical activities of HIV-1 integrase. *Retrovirology* 5, 114.
- Descours, B., Cribier, A., Chable-Bessia, C., Ayinde, D., Rice, G., Crow, Y., Yatim, A., Schawartz, O., Laguette, N., and Benkirane, M. (2012). SAMHD1 restricts HIV-1 reverse transcription in quiescent CD4+ T-cells. *Retrovirology* 9, 87.
- Diaz-Cano, S. (2015). Pathological Bases for a Robust Application of Cancer Molecular Classification. *Int. J. Mol. Sci.* 16, 8655–8675.
- Dick, F. a and Rubin, S.M. (2013). Molecular mechanisms underlying RB protein function. *Nat. Rev. Mol. Cell Biol.* 14, 297–306.
- Dickson, M. a. (2014). Molecular pathways: CDK4 inhibitors for cancer therapy. *Clin. Cancer Res.* 20, 3379–3383.
- Doehle, B.P., Schäfer, A., and Cullen, B.R. (2005). Human APOBEC3B is a potent inhibitor of HIV-1 infectivity and is resistant to HIV-1 Vif. *Virology* 339, 281–288.
- Doehle, B.P., Chang, K., Rustagi, a., McNevin, J., McElrath, M.J., and Gale, M. (2012). Vpu Mediates Depletion of Interferon Regulatory Factor 3 during HIV Infection by a Lysosome-Dependent Mechanism. *J. Virol.* 86, 8367–8374.
- Douglas, J.L., Viswanathan, K., McCarroll, M.N., Gustin, J.K., Früh, K., and Moses, A. V (2009). Vpu directs the degradation of the human immunodeficiency virus restriction factor BST-2/Tetherin via a {beta}TrCP-dependent mechanism. *J. Virol.* 83, 7931–7947.
- Le Du, F., Eckhardt, B.L., Lim, B., Litton, J.K., Meric-bernstam, F., Gonzalez-angulo, A.M., and Naoto, T. (2015). Is the future of personalized therapy in triple-negative breast cancer based on molecular subtype ? *Oncotarget* 6.
- Dubey, S., Satyanarayana, Y.D., and Lavania, H. (2007). Development of integrase inhibitors for treatment of AIDS: An overview. *Eur. J. Med. Chem.* 42, 1159–1168.
- Dunn, J.M., Phillips, R. a, Becker, a J., and Gallie, B.L. (1988). Identification of germline and somatic mutations affecting the retinoblastoma gene. *Science* 241, 1797–1800.
- Dupressoir, a., Vernochet, C., Harper, F., Guegan, J., Dessen, P., Pierron, G., and Heidmann, T. (2011). PNAS Plus: A pair of co-opted retroviral envelope syncytin genes is required for formation of the two-layered murine placental syncytiotrophoblast. *Proc. Natl. Acad. Sci.* 108, E1164–E1173.
- Dupressoir, a., Lavalie, C., and Heidmann, T. (2012). From ancestral infectious retroviruses to bona fide cellular genes: Role of the captured syncytins in placentation. *Placenta* 33, 663–671.
- Dupressoir, A., Marceau, G., Vernochet, C., Bénit, L., Kanellopoulos, C., Sapin, V., and Heidmann, T. (2005). Syncytin-A and syncytin-B, two fusogenic placenta-specific murine envelope genes of retroviral origin conserved in Muridae. *Proc. Natl. Acad. Sci. U. S. A.* 102, 725–730.

- Dupressoir, A., Vernochet, C., Bawa, O., Harper, F., Pierron, G., Opolon, P., and Heidmann, T. (2009). Syncytin-A knockout mice demonstrate the critical role in placentation of a fusogenic, endogenous retrovirus-derived, envelope gene. *Proc. Natl. Acad. Sci. U. S. A.* *106*, 12127–12132.
- Ebrahim, O. and Mazanderani, A.H. (2013). Recent developments in HIV treatment and their dissemination in poor countries. *Infect. Dis. Rep.* *5*, 2–7.
- Ellermann, V. and Bang, O. (1909). Experimentelle Leukämie bei Hühnern. II. Zeitschrift für Hyg. Und Infekt. *63*, 231–272.
- Enquist, L.W. (2002). Exploiting circuit-specific spread of pseudorabies virus in the central nervous system: insights to pathogenesis and circuit tracers. *J. Infect. Dis.* *186 Suppl*, S209–S214.
- Etten, J.L. Van, Lane, L.C., and Dunigan, D.D. (2010). DNA Viruses: The Really Big Ones (Giruses). *Annu Rev Microbiol.* 83–99.
- Fassati, A. (2012). Multiple roles of the capsid protein in the early steps of HIV-1 infection. *Virus Res.* *170*, 15–24.
- Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W., and Lührmann, R. (1995). The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* *82*, 475–483.
- Fishel, R. and Kolodner, R.D. (1995). Identification of mismatch repair genes and their role in the development of cancer. *Curr. Opin. Genet. Dev.* *5*, 382–395.
- Fletcher, T.M., Brichacek, B., Sharova, N., Newman, M. a, Stivahtis, G., Sharp, P.M., Emerman, M., Hahn, B.H., and Stevenson, M. (1996). Nuclear import and cell cycle arrest functions of the HIV-1 Vpr protein are encoded by two separate genes in HIV-2/SIV(SM). *EMBO J.* *15*, 6155–6165.
- Flusberg, D. a. and Sorger, P.K. (2015). Surviving apoptosis: life–death signaling in single cells. *Trends Cell Biol.* 1–13.
- Fornerod, M., Ohno, M., Yoshida, M., and Mattaj, I.W. (1997). CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* *90*, 1051–1060.
- Forterre, P. (2012). The Virocell Concept. *eLS* 1–6.
- Fox, E.J., Prindle, M.J., and Loeb, L.A. (2013). Do mutator mutations fuel tumorigenesis? *Cancer Metastasis* *32*, 353–361.
- Frada, M., Probert, I., Allen, M.J., Wilson, W.H., and de Vargas, C. (2008). The “Cheshire Cat” escape strategy of the coccidiphore *Emiliana huxleyi* in response to viral infection. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 15944–15949.
- Fraenkel-Conrat, H. (1956). The Role of the Nucleic Acid in the Reconstitution of Active Tobacco Mosaic Virus. *J. Am. Chem. Soc.* *78*, 882–883.
- Franzolin, E., Pontarin, G., Rampazzo, C., Miazzi, C., Ferraro, P., Palumbo, E., Reichard, P., and Bianchi, V. (2013). The deoxynucleotide triphosphohydrolase SAMHD1 is a major regulator of DNA precursor pools in mammalian cells. *Proc. Natl. Acad. Sci.* *110*, 14272–14277.

- Friend, S.H., Bernards, R., Rogelj, S., Weinberg, R. a, Rapaport, J.M., Albert, D.M., and Dryja, T.P. (1986). A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* 323, 643–646.
- Galão, R.P., Le Tortorec, A., Pickering, S., Kueck, T., and Neil, S.J.D. (2012). Innate sensing of HIV-1 assembly by tetherin induces NFκB-dependent proinflammatory responses. *Cell Host Microbe* 12, 633–644.
- Gao, D., Wu, J., Wu, Y., Du, F., Aroh, C., Yan, N., Sun, L., and Chen, Z.J. (2013). Cyclic GMP-AMP Synthase Is an Innate Immune Sensor of HIV and Other Retroviruses. 1375, 903–907.
- Gee, C.J. and Harris, H. (1979). Tumorigenicity of Cells Transformed by Simian Virus 40 and of Hybrids Between Such Cells and Normal Diploid Cells. *J. Cell Sci.* 240, 223–240.
- Gerlinger, M., McGranahan, N., Dewhurst, S.M., Burrell, R. a, Tomlinson, I., and Swanton, C. (2014). Cancer: Evolution within a Lifetime. *Annu. Rev. Genet.* 48, 215–238.
- Giono, L.E. and Manfredi, J.J. (2006). The p53 Tumor Suppressor Participates in Multiple Cell Cycle Checkpoints. *J. Cell. Physiol.* 209, 13–20.
- Goffinet, C., Allespach, I., Homann, S., Tervo, H.M., Habermann, A., Rupp, D., Oberbremer, L., Kern, C., Tibroni, N., Welsch, S., Krijnse-Locker, J., Banting, G., Kräusslich, H.G., Fackler, O.T., and Keppler, O.T. (2009). HIV-1 Antagonism of CD317 Is Species Specific and Involves Vpu-Mediated Proteasomal Degradation of the Restriction Factor. *Cell Host Microbe* 5, 285–297.
- Goldstone, D.C., Ennis-Adeniran, V., Hedden, J.J., Groom, H.C.T., Rice, G.I., Christodoulou, E., Walker, P. a., Kelly, G., Haire, L.F., Yap, M.W., de Carvalho, L.P.S., Stoye, J.P., Crow, Y.J., Taylor, I. a., and Webb, M. (2011). HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. *Nature* 480, 379–382.
- Goncalves, A., Karayel, E., Rice, G.I., Bennett, K.L., Crow, Y.J., Superti-Furga, G., and Bürckstümmer, T. (2012). SAMHD1 is a nucleic-acid binding protein that is mislocalized due to aicardi-goutières syndrome-associated mutations. *Hum. Mutat.* 33, 1116–1122.
- Goujon, C., Jarrosson-Wuillème, L., Bernaud, J., Rigal, D., Darlix, J.-L., and Cimorelli, a (2006). With a little help from a friend: increasing HIV transduction of monocyte-derived dendritic cells with virion-like particles of SIV(MAC). *Gene Ther.* 13, 991–994.
- Goujon, C., Rivière, L., Jarrosson-Wuilleme, L., Bernaud, J., Rigal, D., Darlix, J.-L., and Cimorelli, A. (2007). SIVSM/HIV-2 Vpx proteins promote retroviral escape from a proteasome-dependent restriction pathway present in human dendritic cells. *Retrovirology* 4, 2.
- Goujon, C., Arfi, V., Pertel, T., Luban, J., Lienard, J., Rigal, D., Darlix, J.-L., and Cimorelli, A. (2008). Characterization of simian immunodeficiency virus SIVSM/human immunodeficiency virus type 2 Vpx function in human myeloid cells. *J. Virol.* 82, 12335–12345.
- Goyenvallé, A., Vulin, A., Fougereousse, F., Leturcq, F., Kaplan, J.-C., Garcia, L., and Danos, O. (2004). Rescue of dystrophic muscle through U7 snRNA-mediated exon skipping. *Science* 306, 1796–1799.
- Grady, L., Axelrod, D., and Trilling, D. (1970). The SV40 pseudovirus: its potential for general transduction in animal cells. *Proc. Natl. Acad. Sci. U. S. A.* 67, 1886–1893.

- Gramberg, T., Kahle, T., Bloch, N., Wittmann, S., Müllers, E., Daddacha, W., Hofmann, H., Kim, B., Lindemann, D., and Landau, N.R. (2013). Restriction of diverse retroviruses by SAMHD1. *Retrovirology* 10, 26.
- Gray, J.W. and Collins, C. (2000). Genome changes and gene expression in human solid tumors. *Carcinogenesis* 21, 443–452.
- Green, J.C., Course, P. a., and Tarran, G. a. (1996). The life-cycle of *Emiliana huxleyi*: A brief review and a study of relative ploidy levels analysed by flow cytometry. *J. Mar. Syst.* 9, 33–44.
- Greider, C.W. (1990). Telomeres, telomerase and senescence. *Bioessays* 12, 363–369.
- Gringhuis, S.I., Vlist, M. Van Der, Berg, L.M. Van Den, Dunnen, J. Den, Litjens, M., and Geijtenbeek, T.B.H. (2010). HIV-1 exploits innate signaling by TLR8 and DC-SIGN for productive infection of dendritic cells. *Nat. Publ. Gr.* 11, 419–426.
- Groden, J., Thliveris, a, Samowitz, W., Carlson, M., Gelbert, L., Albertsen, H., Joslyn, G., Stevens, J., Spirio, L., and Robertson, M. (1991). Identification and characterization of the familial adenomatous polyposis coli gene. *Cell* 66, 589–600.
- Grüter, P., Tabernero, C., von Kobbe, C., Schmitt, C., Saavedra, C., Bachi, a, Wilm, M., Felber, B.K., and Izaurralde, E. (1998). TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Mol. Cell* 1, 649–659.
- Guerrero, S., Batisse, J., Libre, C., Bernacchi, S., Marquet, R., and Paillart, J.-C. (2015). HIV-1 Replication and the Cellular Eukaryotic Translation Apparatus. *Viruses* 7, 199–218.
- Guo, H., Gao, J., Taxman, D., Ting, J., and Su, L. (2014). HIV-1 Infection Induces Interleukin-1B Production via TLR8 Protein-dependent and NLRP3 Inflammasome Mechanisms in Human Monocytes. 289, 21716–21726.
- Hahn, B.H., Shaw, G.M., De Cock, K.M., and Sharp, P.M. (2000). AIDS as a zoonosis: scientific and public health implications. *Science* 287, 607–614.
- Hanahan, D. and Weinberg, R. a. (2000). The hallmarks of cancer. *Cell* 100, 57–70.
- Hanahan, D. and Weinberg, R. a. (2011). Hallmarks of cancer: The next generation. *Cell* 144, 646–674.
- Hansen, E.C., Seamon, K.J., Cravens, S.L., and Stivers, J.T. (2014). GTP activator and dNTP substrates of HIV-1 restriction factor SAMHD1 generate a long-lived activated state. *Proc. Natl. Acad. Sci. U. S. A.* 111, E1843–E1851.
- Harley, C.B., Futcher, a B., and Greider, C.W. (1990). Telomeres shorten during ageing of human fibroblasts. *Nature* 345, 458–460.
- Harris, R.S., Bishop, K.N., Sheehy, A.M., Craig, H.M., Petersen-Mahrt, S.K., Watt, I.N., Neuberger, M.S., and Malim, M.H. (2003). DNA deamination mediates innate immunity to retroviral infection. *Cell* 113, 803–809.
- Hayflick, L. (1965). the Limited in Vitro Lifetime of Human Diploid Cell Strains. *Exp. Cell Res.* 37, 614–636.

- Hayward, W.S., Neel, B.G., and Astrin, S.M. (1981). Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukemia. *Nature* 290, 475–480.
- He, N., Liu, M., Hsu, J., Xue, Y., Chou, S., Burlingame, A., Krogan, N., Alber, T., and Zhou, Q. (2010). HIV-1 Tat and Host AFF4 Recruit Two Transcription Elongation Factors into a Bifunctional Complex for Coordinated Activation of HIV-1 Transcription. *38*, 428–438.
- Heidmann, O., Vernochet, C., Dupressoir, A., and Heidmann, T. (2009). Identification of an endogenous retroviral envelope gene with fusogenic activity and placenta-specific expression in the rabbit: a new “syncytin” in a third order of mammals. *Retrovirology* 6, 107.
- Heuckmann, J.M. and Thomas, R.K. (2015). A new generation of cancer genome diagnostics for routine clinical use: overcoming the roadblocks to personalized cancer medicine. *Ann. Oncol.* 1–8.
- Hirsch, V.M., Sharkey, M.E., Brown, C.R., Brichacek, B., Goldstein, S., Wakefield, J., Byrum, R., Elkins, W.R., Hahn, B.H., Lifson, J.D., and Stevenson, M. (1998). Vpx is required for dissemination and pathogenesis of SIV(SM) PBj: evidence of macrophage-dependent viral amplification. *Nat. Med.* 4, 1401–1408.
- Hollenbaugh, J. a., Gee, P., Baker, J., Daly, M.B., Amie, S.M., Tate, J., Kasai, N., Kanemura, Y., Kim, D.H., Ward, B.M., Koyanagi, Y., and Kim, B. (2013). Host Factor SAMHD1 Restricts DNA Viruses in Non-Dividing Myeloid Cells. *PLoS Pathog.* 9.
- Hollis, G.F., Hieter, P. a, McBride, O.W., Swan, D., and Leder, P. (1982). Processed genes: a dispersed human immunoglobulin gene bearing evidence of RNA-type processing. *Nature* 296, 321–325.
- Hotter, D., Kirchhoff, F., and Sauter, D. (2013). HIV-1 Vpu does not degrade interferon regulatory factor 3. *J. Virol.* 87, 7160–7165.
- Houdan, A., Billard, C., Marie, D., Not, F., Sáez, A.G., Young, J.R., and Probert, I. (2004). Holococcolithophore-heterococcolithophore (Haptophyta) life cycles: Flow cytometric analysis of relative ploidy levels. *Syst. Biodivers.* 1, 453–465.
- Hrecka, K., Hao, C., Gierszewska, M., Swanson, S.K., Kesik-Brodacka, M., Srivastava, S., Florens, L., Washburn, M.P., and Skowronski, J. (2011). Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature* 474, 658–661.
- Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M.J., Kenzelmann-Broz, D., Khalil, A.M., Zuk, O., Amit, I., Rabani, M., Attardi, L.D., Regev, A., Lander, E.S., Jacks, T., and Rinn, J.L. (2010). A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* 142, 409–419.
- Huasong, L., Zichong, L., Yuhua, X., and Zhou, Q. (2013). Viral-host interactions that control HIV-1 transcriptional elongation. *Chem Rev.* 29, 997–1003.
- Hung, T., Wang, Y., Lin, M.F., Koegel, A.K., Kotake, Y., Grant, G.D., Horlings, H.M., Shah, N., Umbricht, C., Wang, P., Wang, Y., Kong, B., Langerød, A., Børresen-Dale, A.-L., Kim, S.K., van de Vijver, M., Sukumar, S., Whitfield, M.L., Kellis, M., Xiong, Y., Wong, D.J., and Chang, H.Y. (2011). Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat. Genet.* 43, 621–629.

Hutten, S. and Kehlenbach, R.H. (2007). CRM1-mediated nuclear export: to the pore and beyond. *Trends Cell Biol.* 17, 193–201.

Iglesias-Rodriguez, M.D., Halloran, P.R., Rickaby, R.E.M., Hall, I.R., Colmenero-Hidalgo, E., Gittins, J.R., Green, D.R.H., Tyrrell, T., Gibbs, S.J., von Dassow, P., Rehm, E., Armbrust, E.V., and Boessenkool, K.P. (2008). Phytoplankton Calcification in a High-CO₂ World. *Science* (80-.). 320, 336–340.

Islam, F., Gopalan, V., Smith, R. a., and Lam, A.K.-Y. (2015). Translational potential of cancer stem cells: A review of the detection of cancer stem cells and their roles in cancer recurrence and cancer treatment. *Exp. Cell Res.* 335, 135–147.

Iwabu, Y., Fujita, H., Kinomoto, M., Kaneko, K., Ishizaka, Y., Tanaka, Y., Sata, T., and Tokunaga, K. (2009). HIV-1 accessory protein Vpu internalizes cell-surface BST-2/tetherin through transmembrane interactions leading to lysosomes. *J. Biol. Chem.* 284, 35060–35072.

Jäger, S., Kim, D.Y., Hultquist, J.F., Shindo, K., Larue, R., Kwon, E., Li, M., Anderson, B., Yen, L., Stanley, S., Mahon, C., Kane, J., Franks-skiba, K., Cimerancic, P., Burlingame, A., Sali, A., Craik, C., Harris, R., Gross, J., and Krogan, N. (2011). Vif hijacks CBF – β to degrade APOBEC3G and promote HIV–1 infection. *Nature* 481, 371–375.

Jakobsen, M.R., Bak, R.O., Andersen, A., Berg, R.K., Jensen, S.B., Jin, T., Laustsen, A., Hansen, K., Østergaard, L., Fitzgerald, K. a, Xiao, T.S., Mikkelsen, J.G., Mogensen, T.H., and Paludan, S.R. (2013). IFI16 senses DNA forms of the lentiviral replication cycle and controls HIV-1 replication. *Proc. Natl. Acad. Sci.* 110, 19651–19651.

Ji, X., Wu, Y., Yan, J., Mehrens, J., Yang, H., DeLucia, M., Hao, C., Gronenborn, A.M., Skowronski, J., Ahn, J., and Xiong, Y. (2013). Mechanism of allosteric activation of SAMHD1 by dGTP. *Nat. Struct. Mol. Biol.* 20, 1304–1309.

Ji, X., Tang, C., Zhao, Q., Wang, W., and Xiong, Y. (2014). Structural basis of cellular dNTP regulation by SAMHD1. *Proc. Natl. Acad. Sci.* 111, E4305–E4314.

Jin, C., Peng, X., Liu, F., Cheng, L., Lu, X., Yao, H., Wu, H., and Wu, N. (2014). MicroRNA-181 expression regulates specific post-transcriptional level of SAMHD1 expression in vitro. *Biochem. Biophys. Res. Commun.* 452, 760–767.

Johnson, D.G. and Walker, C.L. (1999). Cyclins and cell cycle checkpoints. *Annu. Rev. Pharmacol. Toxicol.* 39, 295–312.

Jones, P. a and Baylin, S.B. (2002). The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* 3, 415–428.

Jouvenet, N., Neil, S.J.D., Zhadina, M., Zang, T., Kratovac, Z., Lee, Y., McNatt, M., Hatzioannou, T., and Bieniasz, P.D. (2009). Broad-spectrum inhibition of retroviral and filoviral particle release by tetherin. *J. Virol.* 83, 1837–1844.

JugeMorineau, N., Harousseau, J.L., Amiot, M., and Bataille, R. (1997). The retinoblastoma susceptibility gene RB-1 in multiple myeloma. *Leuk. Lymphoma* 24, 229–237.

Kasher, P.R., Jenkinson, E.M., Briolat, V., Gent, D., Morrissey, C., Zeef, L. a. H., Rice, G.I., Levraud, J.-P., and Crow, Y.J. (2015). Characterization of samhd1 Morphant Zebrafish Recapitulates Features of the Human Type I Interferonopathy Aicardi-Goutieres Syndrome. *J. Immunol.* 194, 2819–2825.

Kaushik, R., Zhu, X., Stranska, R., Wu, Y., and Stevenson, M. (2009). A Cellular Restriction Dictates the Permissivity of Nondividing Monocytes/Macrophages to Lentivirus and Gammaretrovirus Infection. *Cell Host Microbe* 6, 68–80.

Kim, C. a and Bowie, J.U. (2003). SAM domains: uniform structure, diversity of function. *Trends Biochem. Sci.* 28, 625–628.

Kim, Y.-J. and Wilson III, D.M. (2011). Overview of Base Excision Repair Biochemistry. *Curr. Mol. Pharmacol.* 5, 3–13.

Kim, B., Nguyen, L. a., Daddacha, W., and Hollenbaugh, J. a. (2012). Tight interplay among SAMHD1 protein level, cellular dNTP levels, and HIV-1 proviral DNA synthesis kinetics in human primary monocyte-derived macrophages. *J. Biol. Chem.* 287, 21570–21574.

Kim, E.T., White, T.E., Brandariz-Núñez, A., Diaz-Griffero, F., and Weitzman, M.D. (2013). SAMHD1 restricts herpes simplex virus 1 in macrophages by limiting DNA replication. *J. Virol.* 87, 12949–12956.

Kinzler, K.W., Nilbert, M.C., Su, L.K., Vogelstein, B., Bryan, T.M., Levy, D.B., Smith, K.J., Preisinger, a C., Hedge, P., and McKechnie, D. (1991). Identification of FAP locus genes from chromosome 5q21. *Science* 253, 661–665.

Klarmann, G.J., Chen, X., North, T.W., and Preston, B.D. (2003). Incorporation of uracil into minus strand DNA affects the specificity of plus strand synthesis initiation during lentiviral reverse transcription. *J. Biol. Chem.* 278, 7902–7909.

Klasse, P. (2012). The molecular basis of HIV entry. *Cell Microbiology* 29, 997–1003.

Klein, G. (1988). Oncogenes and tumor suppressor genes. *Acta Oncol.* 27, 427–437.

Knauer, S.K., Bier, C., Habtemichael, N., and Stauber, R.H. (2006). The Survivin-Crm1 interaction is essential for chromosomal passenger complex localization and function. *EMBO Rep.* 7, 1259–1265.

Koharudin, L., Wu, Y., DeLucia, M., Mehrens, J., Gronenborn, A., and Ahn, J. (2014). Structural basis of allosteric activation of sterile α motif and histidine-aspartate domain-containing protein 1 (SAMHD1) by nucleoside triphosphates. Koharudin 2014.pdf. *J. Biol. Chem.* 289, 32617–32627.

Kojima, K., Kornblau, S.M., Ruvolo, V., Dilip, A., Duvvuri, S., Davis, R.E., Zhang, M., Wang, Z., Coombes, K.R., Zhang, N., Qiu, Y.H., Burks, J.K., Kantarjian, H., Shacham, S., Kauffman, M., Andreeff, M., and Dc, W. (2013). Prognostic impact and targeting of CRM1 in acute myeloid leukemia. *Blood* 121, 4166–4174.

Kong, R., Li, H., Bibollet-Ruche, F., Decker, J.M., Zheng, N.N., Gottlieb, G.S., Kiviat, N.B., Sow, P.S., Georgiev, I., Hahn, B.H., Kwong, P.D., Robinson, J.E., and Shaw, G.M. (2012). Broad and Potent Neutralizing Antibody Responses Elicited in Natural HIV-2 Infection. *J. Virol.* 86, 947–960.

Korsmeyer, S.J., Wei, M.C., Saito, M., Weiler, S., Oh, K.J., and Schlesinger, P.H. (2000). Pro-apoptotic cascade activates BID, which oligomerizes BAK or BAX into pores that result in the release of cytochrome c. *Cell Death Differ.* 7, 1166–1173.

Kretschmer, S., Wolf, C., König, N., Staroske, W., Guck, J., Häusler, M., Luksch, H., Nguyen, L. a, Kim, B., Alexopoulou, D., Dahl, A., Rapp, A., Cardoso, M.C., Shevchenko, A., and Lee-Kirsch, M.A. (2014). SAMHD1 prevents autoimmunity by maintaining genome stability. *Ann. Rheum. Dis.* 0, 1–8.

Kröger, A., Köster, M., Schroeder, K., Hauser, H., and Mueller, P.P. (2002). Activities of IRF-1. *J. Interferon Cytokine Res.* 22, 5–14.

Kvaratskhelia, M., Sharma, A., Larue, R.C., Serrao, E., and Engelman, A. (2014). Molecular mechanisms of retroviral integration site selection. *Nucleic Acids Res.* 42, gku769 – .

Kyei, G.B., Cheng, X., Ramani, R., and Ratner, L. (2015). Article Cyclin L2 Is a Critical HIV Dependency Factor in Macrophages that Controls SAMHD1 Abundance. *Cell Host Microbe* 17, 98–106.

Lafon, M. (2005). Rabies virus receptors. *J. Neurovirol.* 11, 82–87.

Laguette, N., Sobhian, B., Casartelli, N., Ringeard, M., Chable-Bessia, C., Ségéral, E., Yatim, A., Emiliani, S., Schwartz, O., and Benkirane, M. (2011). SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* 474, 654–657.

Lahaye, X., Satoh, T., Gentili, M., Cerboni, S., Conrad, C., Hurbain, I., Marjou, A. El, Lacabaratz, C., Lelièvre, J., and Manel, N. (2013). The Capsids of HIV-1 and HIV-2 Determine Immune Detection of the Viral cDNA by the Innate Sensor cGAS in Dendritic Cells. 1132–1142.

Lahouassa, H., Daddacha, W., Hofmann, H., Ayinde, D., Logue, E.C., Dragin, L., Bloch, N., Maudet, C., Bertrand, M., Gramberg, T., Pancino, G., Priet, S., Canard, B., Laguette, N., Benkirane, M., Transy, C., Landau, N.R., Kim, B., and Margottin-Goguet, F. (2012). SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nat. Immunol.* 13, 621–621.

Lautenberger, J. a, Schulz, R. a, Garon, C.F., Tschlis, P.N., Spyropoulos, D.D., Pry, T.W., Rushlow, K.E., and Papas, T.S. (1981). The transforming sequences of avian myelocytomatosis virus (MC29). *J. Supramol. Struct. Cell. Biochem.* 16, 193–207.

Lecossier, D., Bouchonnet, F., Clavel, F., and Hance, A.J. (2003). Hypermutation of HIV-1 DNA in the absence of the Vif protein. *Science* 300, 1112.

Leek, R.D., Kaklamanis, L., Pezzella, F., Gatter, K.C., and Harris, a L. (1994). Bcl-2 in Normal Human Breast and Carcinoma, Association With Oestrogen Receptor-Positive, Epidermal Growth Factor Receptor-Negative Tumours and in Situ Cancer. *Br. J. Cancer* 69, 135–139.

Lepelley, A., Louis, S., Sourisseau, M., Law, H.K.W., Pothlichet, J., Schilte, C., Chaperot, L., Plumas, J., Randall, R.E., Si-tahar, M., Mammano, F., Albert, M.L., and Schwartz, O. (2011). Innate Sensing of HIV-Infected Cells. 7.

Leshinsky-Silver, E., Maling, G., Ben-Sira, L., Kidron, D., Cohen, S., Inbar, S., Bezaleli, T., Levine, A., Vinkler, C., Lev, D., and Lerman-Sagie, T. (2011). A large homozygous deletion in the SAMHD1 gene

causes atypical Aicardi-Goutières syndrome associated with mtDNA deletions. *Eur. J. Hum. Genet.* 19, 287–292.

Lévesque, L., Guzik, B., Guan, T., Coyle, J., Black, B.E., Rekosh, D., Hammarskjöld, M.L., and Paschal, B.M. (2001). RNA Export Mediated by Tap Involves NXT1-dependent Interactions with the Nuclear Pore Complex. *J. Biol. Chem.* 276, 44953–44962.

Levine, a J. (1993). The tumor suppressor genes. *Annu. Rev. Biochem.* 62, 623–651.

Levine, a J. and Teresky, a K. (1970). Deoxyribonucleic acid replication in simian virus 40-infected cells. II. Detection and characterization of simian virus 40 pseudovirions. *J. Virol.* 5, 451–457.

Li, J., Poi, M.J., and Tsai, M.-D. (2011). The Regulatory Mechanisms of Tumor Suppressor p16INK4 and Relevance to cancer. *Biochemistry* 50, 5566–5582.

Li, N., Zhang, W., and Cao, X. (2000). Identification of human homologue of mouse IFN- γ induced protein from human dendritic cells. *Immunol. Lett.* 74, 221–224.

Li, Z., Owonikoko, T.K., Sun, S.-Y., Ramalingam, S.S., Doetsch, P.W., Xiao, Z.-Q., Khuri, F.R., Curran, W.J., and Deng, X. (2012). c-Myc Suppression of DNA Double-strand Break Repair. *Neoplasia* 14, 1190–IN35.

Liao, W., Bao, Z., Cheng, C., Mok, Y.-K., and Wong, W.S.F. (2008). Dendritic cell-derived interferon-gamma-induced protein mediates tumor necrosis factor-alpha stimulation of human lung fibroblasts. *Proteomics* 8, 2640–2650.

Liebermann, D.A.N., Hoffman-liebermann, B., Troutt, A.B., Kedes, L., and Cohen, S.N. (1986). Sequences from Sea Urchin TU Transposons Are Conserved among Multiple Eucaryotic Species, Including Humans. *Mol. Cell. Biol.* 6, 218–226.

Lienlaf, M., Hayashi, F., Di Nunzio, F., Tochio, N., Kigawa, T., Yokoyama, S., and Diaz-Griffero, F. (2011). Contribution of E3-Ubiquitin Ligase Activity to HIV-1 Restriction by TRIM5 α : Structure of the RING Domain of TRIM5 α . *J. Virol.* 85, 8725–8737.

Van Lint, C., Bouchat, S., and Marcello, A. (2013). HIV-1 transcription and latency: an update. *Retrovirology* 10, 67.

Liu, M. a. (2010). Immunologic basis of vaccine vectors. *Immunity* 33, 504–515.

Liu, Y. and Lu, X. (2012). Non-coding RNAs in DNA damage response. *Am. J. Cancer Res.* 2, 658–675.

Liu, B., Sarkis, P.T.N., Luo, K., Yu, Y., and Yu, X.-F. (2005). Regulation of Apobec3F and Human Immunodeficiency Virus Type 1 Vif by Regulation of Apobec3F and Human Immunodeficiency Virus Type 1 Vif by Vif-Cul5-ElonB / C E3 Ubiquitin Ligase. *J. Virol.* 79, 9579–9587.

Lively, C.M. and Morran, L.T. (2014). The ecology of sexual reproduction. *J. Evol. Biol.* 27, 1292–1303.

Loewy, a. D. (1998). Viruses as transneuronal tracers for defining neural circuits. *Neurosci. Biobehav. Rev.* 22, 679–684.

Loisel-Meyer, S., Swainson, L., Craveiro, M., Oburoglu, L., Mongellaz, C., Costa, C., Martinez, M., Cosset, F.-L., Battini, J.-L., Herzenberg, L. a., Herzenberg, L. a., Atkuri, K.R., Sitbon, M., Kinet, S., Verhoeven, E., and Taylor, N. (2012). Glut1-mediated glucose transport regulates HIV infection. *Proc. Natl. Acad. Sci.* 109, 2549–2554.

Lorin, C., Combredet, C., Labrousse, V., Mollet, L., Desprès, P., and Tangy, F. (2005). A paediatric vaccination vector based on live attenuated measles vaccine. *Therapie* 60, 227–233.

Lu, Q.L., Elia, G., Lucas, S., and Thomas, J.A. (1993). Bcl-2 proto-oncogene expression in Epstein-Barr-virus-associated nasopharyngeal carcinoma. *Int. J. Cancer* 53, 29–35.

Lueders, K.K. and Kuff, E.L. (1977). Sequences associated with intracisternal A particles are reiterated in the mouse genome. *Cell* 12, 963–972.

Luo, J., Luo, Y., Sun, J., Zhou, Y., Zhang, Y., and Yang, X. (2015). Adeno-associated virus-mediated cancer gene therapy: Current status. *Cancer Lett.* 356, 347–356.

Macpherson, I. (1970). The characteristics of animal cells transformed in vitro. *Adv. Cancer Res.* 13, 169–215.

Madani, N. and Kabat, D. (1998). An endogenous inhibitor of human immunodeficiency virus in human lymphocytes is overcome by the viral Vif protein. *J. Virol.* 72, 10251–10255.

Manel, N., Hogstad, B., Wang, Y., Levy, D.E., Unutmaz, D., and Littman, D.R. (2010). A cryptic sensor for HIV-1 activates antiviral innate immunity in dendritic cells. *Nature* 467, 214–217.

Mangeat, B., Gers-Huber, G., Lehmann, M., Zufferey, M., Luban, J., and Piguet, V. (2009). HIV-1 Vpu neutralizes the antiviral factor tetherin/BST-2 by binding it and directing its beta-TrCP2-dependent degradation. *PLoS Pathog.* 5.

Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L., and Trono, D. (2003). Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* 424, 99–103.

Manno, C.S., Pierce, G.F., Arruda, V.R., Glader, B., Ragni, M., Rasko, J.J., Ozelo, M.C., Hoots, K., Blatt, P., Konkle, B., Dake, M., Kaye, R., Razavi, M., Zajko, A., Zehnder, J., Rustagi, P.K., Nakai, H., Chew, A., Leonard, D., Wright, J.F., Lessard, R.R., Sommer, J.M., Tigges, M., Sabatino, D., Luk, A., Jiang, H., Mingozi, F., Couto, L., Ertl, H.C., High, K. a, and Kay, M. a (2006). Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nat. Med.* 12, 342–347.

Margottin, F., Bour, S.P., Durand, H., Selig, L., Benichou, S., Richard, V., Thomas, D., Strebel, K., and Benarous, R. (1998). A novel human WD protein, h-beta TrCp, that interacts with HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif. *Mol. Cell* 1, 565–574.

Mariani, R., Chen, D., Schröfelbauer, B., Navarro, F., König, R., Bollman, B., Münk, C., Nymark-McMahon, H., and Landau, N.R. (2003). Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell* 114, 21–31.

Marlink, R., Kanki, P., Thior, I., Travers, K., Eisen, G., Siby, T., Traore, I., Hsieh, C.C., Dia, M.C., and Gueye, E.H. (1994). Reduced rate of disease development after HIV-2 infection as compared to HIV-1. *Science* 265, 1587–1590.

Martin, X. and Dolivo, M. (1983). Neuronal and transneuronal tracing in the trigeminal system of the rat using the herpes virus suis. *Brain Res.* 273, 253–276.

Martínez Martínez, J., Schroeder, D.C., Larsen, A., Bratbak, G., and Wilson, W.H. (2007). Molecular dynamics of *Emiliana huxleyi* and cooccurring viruses during two separate mesocosm studies. *Appl. Environ. Microbiol.* 73, 554–562.

Mátrai, J., Chuah, M.K.L., and VandenDriessche, T. (2010). Recent advances in lentiviral vector development and applications. *Mol. Ther.* 18, 477–490.

Matreyek, K. a. and Engelman, A. (2013). Viral and cellular requirements for the nuclear entry of retroviral preintegration nucleoprotein complexes. *Viruses* 5, 2483–2511.

McClintock, B. (1951). CHROMOSOME ORGANIZATION AND GENIC EXPRESSION. COLD SPRING Harb. Symp. Quant. Biol. 16, 13–47.

McGeoch, D.J., Dalrymple, M. a., Davison, a. J., Dolan, a., Frame, M.C., McNab, D., Perry, L.J., Scott, J.E., and Taylor, P. (1988). The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. *J. Gen. Virol.* 69, 1531–1574.

Meacham, C.E. and Morrison, S.J. (2013). Tumour heterogeneity and cancer cell plasticity. *Nature* 501, 328–337.

Meier, A., Alter, G., Frahm, N., Sidhu, H., Li, B., Bagchi, A., Teigen, N., Streeck, H., Stellbrink, H., Hellman, J., Lunzen, J. Van, and Altfeld, M. (2007). MyD88-Dependent Immune Activation Mediated by Human Immunodeficiency Virus Type 1-Encoded Toll-Like Receptors Ligands. 81, 8180–8191.

Mi, S., Lee, X., Li, X., Veldman, G.M., Finnerty, H., Racie, L., LaVallie, E., Tang, X.Y., Edouard, P., Howes, S., Keith, J.C., and McCoy, J.M. (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature* 403, 785–789.

Miazzi, C., Ferraro, P., Pontarin, G., Rampazzo, C., Reichard, P., and Bianchi, V. (2014). Allosteric regulation of the human and mouse deoxyribonucleotide triphosphohydrolase sterile α -motif /histidine-aspartate domain containing protein 1 (SAMHD1). *J. Biol. Chem.*

Michael H. Malim, Joachim Hauber, Shu-Yun Le, J.V.M.& B.R.C. (1989). The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* 338.

Michalek, R.D. and Rathmell, J.C. (2011). The metabolic life and times of a T cell. *Immunol Rev* 236, 190–202.

Miki, Y., Swensen, J., Shattuck-eidens, D., Futreal, P.A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L.M., Ding, W., Bell, R., Rosenthal, J., Hussey, C., Tran, T., McClure, M., Frye, C., Hattier, T., Phelps, R., Haugen-strano, A., Katcher, H., Yakumo, K., Gholami, Z., Shaffer, D., Stone, S., Bayer, S., Wray, C., Bogden, R., Dayananth, P., Ward, J., Tonin, P., Narod, S., Bristow, P.K., Norris, F.H., Helvering, L., Morrison, P., Rosteck, P., Lai, M., Barrett, J.C., Lewis, C., Neuhausen, S., Cannon-

- albright, L., Goldgar, D., Wiseman, R., Kamb, A., and Skolnick, M.H. (1994). Strong Candidate for the Breast and Ovarian Cancer Susceptibility Gene BRCA1. *Science* (80-.). 266, 66–71.
- Van Miltenburg, M.H. and Jonkers, J. (2012). Using genetically engineered mouse models to validate candidate cancer genes and test new therapeutic approaches. *Curr. Opin. Genet. Dev.* 22, 21–27.
- Mitchell, R.S., Katsura, C., Skasko, M. a., Fitzpatrick, K., Lau, D., Ruiz, A., Stephens, E.B., Margottin-Goguet, F., Benarous, R., and Guatelli, J.C. (2009). Vpu antagonizes BST-2-mediated restriction of HIV-1 release via β -TrCP and endo-lysosomal trafficking. *PLoS Pathog.* 5.
- Miyagi, E., Andrew, A.J., Kao, S., and Strebel, K. (2009). Vpu enhances HIV-1 virus release in the absence of Bst-2 cell surface down-modulation and intracellular depletion. *Proc. Natl. Acad. Sci. U. S. A.* 106, 2868–2873.
- Molinari, M. (2000). Cell cycle checkpoints and their inactivation in human cancer. *Cell Prolif* 33, 261–274.
- Morgan, D.O. (1995). Principles of CDK regulation. *Nature* 374, 131–134.
- Moroianu, J. and Blobel, G. (1995). Protein export from the nucleus requires the GTPase Ran and GTP hydrolysis. *Proc. Natl. Acad. Sci. U. S. A.* 92, 4318–4322.
- Moroianu, J., Blobel, G., and Radu, a (1995). Previously identified protein of uncertain function is karyopherin alpha and together with karyopherin beta docks import substrate at nuclear pore complexes. *Proc. Natl. Acad. Sci. U. S. A.* 92, 2008–2011.
- Morrissey, C., Schwefel, D., Ennis-Adeniran, V., Taylor, I. a., Crow, Y.J., and Webb, M. (2015). The eukaryotic elongation factor eEF1A1 interacts with SAMHD1. *Biochem. J.* 466, 69–76.
- Murray, M.J., Shilo, B.Z., Shih, C., Cowing, D., Hsu, H.W., and Weinberg, R. a (1981). Three different human tumor cell lines contain different oncogenes. *Cell* 25, 355–361.
- N. D. Zinder and J. Lederberg (1952). Genetic exchange in Salmonella. *J. Bacteriol.* 679–699.
- Nakamura, Y. (analytic) (1995). The adenomatous polyposis coli gene and human cancers. *J. Cancer Res. Clin. Oncol.* 121, 529–534.
- Naldini, L., Blömer, U., Gally, P., Ory, D., Mulligan, R., Gage, F.H., Verma, I.M., and Trono, D. (1996). In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* 272, 263–267.
- Narkpuk, J., Wanitchang, A., Kramyu, J., Frantz, P.N., Jongkaewwattana, A., and Teeravechyan, S. (2014). An unconventional BST-2 function: Down-regulation of transient protein expression. *Biochem. Biophys. Res. Commun.* 450, 1469–1474.
- Neil, S.J.D., Zang, T., and Bieniasz, P.D. (2008). Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* 451, 425–430.
- Neville, M., Stutz, F., Lee, L., Davis, L.I., and Rosbash, M. (1997). The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Curr. Biol.* 7, 767–775.

- Nishioka, Y., Leder, a, and Leder, P. (1980). Unusual alpha-globin-like gene that has cleanly lost both globin intervening sequences. *Proc. Natl. Acad. Sci. U. S. A.* 77, 2806–2809.
- Nzonza, A., Lecollinet, S., Chat, S., Lowenski, S., Mérour, E., Biacchesi, S., and Brémont, M. (2014). A recombinant Novirhabdovirus presenting at the surface the e glycoprotein from West Nile Virus (WNV) is immunogenic and provides partial protection against lethal WNV challenge in BALB/c mice. *PLoS One* 9.
- Okumura, A., Alce, T., Lubyova, B., Ezelle, H., Strebel, K., and Pitha, P. (2008). HIV-1 accessory proteins VPR and Vif modulate antiviral response by targeting IRF-3 for degradation. *Virology* 373, 85–97.
- Olubajo, B., Mitchell-Fearon, K., and Ogunmoroti, O. (2014). A comparative systematic review of the optimal CD4 cell count threshold for HIV treatment initiation. *Interdiscip. Perspect. Infect. Dis.* 2014, 7–11.
- Onder, Z., Chang, V., and Moroianu, J. (2015). Nuclear export of cutaneous HPV8 E7 oncoprotein is mediated by a leucine-rich nuclear export signal via a CRM1 pathway. *Virology* 474, 28–33.
- Ono, A. and Freed, E.O. (2001). Plasma membrane rafts play a critical role in HIV-1 assembly and release. *Proc. Natl. Acad. Sci. U. S. A.* 98, 13925–13930.
- Paca, R.E., Ogert, R. a, Hibbert, C.S., Izaurralde, E., and Beemon, K.L. (2000). Rous sarcoma virus DR posttranscriptional elements use a novel RNA export pathway. *J. Virol.* 74, 9507–9514.
- Pacheco, B., Finzi, A., Stremlau, M., and Sodroski, J. (2010). Adaptation of HIV-1 to Cells Expressing Rhesus Monkey TRIM5 α . *Virology*.
- Palazzo, A.F. and Gregory, T.R. (2014). The Case for Junk DNA. *PLoS Genet.* 10.
- Palmer, C.S. and Crowe, S.M. (2012). The role of glucose and lipid metabolism in the pathogenesis of HIV-1 infection. *Curr. Trends Immunol.* 13, 37–50.
- Palmer, C.S., Ostrowski, M., Balderson, B., Christian, N., and Crowe, S.M. (2015). Glucose Metabolism Regulates T Cell Activation, Differentiation, and Functions. *Front. Immunol.* 6, 1–6.
- Park, S.Y., Waheed, a. a., Zhang, Z.-R., Freed, E.O., and Bonifacino, J.S. (2014). HIV-1 Vpu Accessory Protein Induces Caspase-mediated Cleavage of IRF3 Transcription Factor. *J. Biol. Chem.* 289, 35102–35110.
- Parrish, A.B., Freel, C.D., and Kornbluth, S. (2013). Cellular Mechanisms Controlling Caspase Activation and Function. *Cold Spring Harb. Perspect. Biol.* 1–24.
- De Parseval, N., Diop, G., Blaise, S., Helle, F., Vasilescu, A., Matsuda, F., and Heidmann, T. (2005). Comprehensive search for intra- and inter-specific sequence polymorphisms among coding envelope genes of retroviral origin found in the human genome: genes and pseudogenes. *BMC Genomics* 6, 117.
- Pauls, E., Jimenez, E., Ruiz, A., Permanyer, M., Ballana, E., Costa, H., Nascimiento, R., Parkhouse, R.M., Peña, R., Riveiro-Muñoz, E., Martinez, M. a, Clotet, B., Esté, J. a, and Bofill, M. (2013).

Restriction of HIV-1 replication in primary macrophages by IL-12 and IL-18 through the upregulation of SAMHD1. *J. Immunol.* **190**, 4736–4741.

Pauls, E., Ruiz, A., Riveira-Muñoz, E., Permanyer, M., Badia, R., Clotet, B., Keppler, O.T., Ballana, E., and Este, J. a (2014a). p21 regulates the HIV-1 restriction factor SAMHD1. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E1322–E1324.

Pauls, E., Ruiz, A., Badia, R., Permanyer, M., Gubern, A., Riveira-Muñoz, E., Torres-Torronteras, J., Alvarez, M., Mothe, B., Brander, C., Crespo, M., Menéndez-Arias, L., Clotet, B., Keppler, O.T., Martí, R., Posas, F., Ballana, E., and Esté, J. a (2014b). Cell Cycle Control and HIV-1 Susceptibility Are Linked by CDK6-Dependent CDK2 Phosphorylation of SAMHD1 in Myeloid and Lymphoid Cells. *J. Immunol.* **193**, 1988–1997.

Pearson, S., Jia, H., and Kandachi, K. (2004). China approves first gene therapy. *Nat. Biotechnol.* **22**, 3–4.

Perez-caballero, D., Zang, T., Ebrahimi, A., McNatt, M.W., Gregory, D. a, Johnson, M.C., and Bieniasz, P.D. (2009). Tetherin inhibits HIV-1 release by directly tethering virions to cells. *Cell* **139**, 499–511.

Pertel, T., Hausmann, S., Morger, D., Züger, S., Guerra, J., Lascano, J., Reinhard, C., Santoni, F., Uchil, P.D., Bisiaux, A., Albert, M., Strambio-de-castillia, C., Pizzato, M., Grütter, M., and Luban, J. (2011). TRIM5 is an innate immune sensor for the retrovirus capsid lattice. *Nature* **472**, 361–365.

Pietras, K. and Östman, A. (2010). Hallmarks of cancer: Interactions with the tumor stroma. *Exp. Cell Res.* **316**, 1324–1331.

Poli, J., Tsaponina, O., Crabbé, L., Keszthelyi, A., Pantesco, V., Chabes, A., Lengronne, A., and Pasero, P. (2012). dNTP pools determine fork progression and origin usage under replication stress. *EMBO J.* **31**, 883–894.

Pothof, J., Verkaik, N.S., van IJcken, W., Wiemer, E. a C., Ta, V.T.B., van der Horst, G.T.J., Jaspers, N.G.J., van Gent, D.C., Hoeijmakers, J.H.J., and Persengiev, S.P. (2009). MicroRNA-mediated gene silencing modulates the UV-induced DNA-damage response. *EMBO J.* **28**, 2090–2099.

Powell, R.D., Holland, P.J., Hollis, T., and Perrino, F.W. (2011). Aicardi-Goutières syndrome gene and HIV-1 restriction factor SAMHD1 is a dGTP-regulated deoxynucleotide triphosphohydrolase. *J. Biol. Chem.* **286**, 43596–43600.

Prakash, R., Zhang, Y., Feng, W., and Jasin, M. (2015). Homologous Recombination and Human Health: The Roles of BRCA1, BRCA2, and Associated Proteins. *Cold Spring Harb. Perspect. Biol.* **7**.

Qasba, P.K. and Aposhian, H. V (1971). DNA and gene therapy: transfer of mouse DNA to human and mouse embryonic cells by polyoma pseudovirions. *Proc. Natl. Acad. Sci. U. S. A.* **68**, 2345–2349.

Ramesh, V., Bernardi, B., Stafa, A., Garone, C., Franzoni, E., Abinun, M., Mitchell, P., Mitra, D., Friswell, M., Nelson, J., Shalev, S. a., Rice, G.I., Gornall, H., Szykiewicz, M., Aymard, F., Ganesan, V., Prendiville, J., Livingston, J.H., and Crow, Y.J. (2010). Intracerebral large artery disease in Aicardi-Goutières syndrome implicates SAMHD1 in vascular homeostasis. *Dev. Med. Child Neurol.* **52**, 725–732.

- Raper, S.E., Chirmule, N., Lee, F.S., Wivel, N. a., Bagg, A., Gao, G.P., Wilson, J.M., and Batshaw, M.L. (2003). Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. *Mol. Genet. Metab.* *80*, 148–158.
- Rasheed, S., Yan, J.S., Lau, A., and Chan, A.S. (2008). HIV replication enhances production of free fatty acids, low density lipoproteins and many key proteins involved in lipid metabolism: A proteomics study. *PLoS One* *3*.
- Rassool, F.V. (2003). DNA double strand breaks (DSB) and non-homologous end joining (NHEJ) pathways in human leukemia. *Cancer Lett.* *193*, 1–9.
- Ratner, L., Gallo, R.C., and Wong-Staal, F. (1985). HTLV-III, LAV, ARV are variants of same AIDS virus. *Nature* *313*.
- Ravenscroft, J.C., Suri, M., Rice, G.I., Szykiewicz, M., and Crow, Y.J. (2011). Autosomal dominant inheritance of a heterozygous mutation in SAMHD1 causing familial chilblain lupus. *Am. J. Med. Genet. Part A* *155*, 235–237.
- Rayess, H., Wang, M.B., and Srivatsan, E.S. (2012). Cellular senescence and tumor suppressor gene p16. *Int. J. Cancer* *130*, 1715–1725.
- Rehwinkel, J. (2014). Mouse knockout models for HIV-1 restriction factors. *Cell. Mol. Life Sci.* *374*–376.
- Rehwinkel, J., Maelfait, J., Bridgeman, A., Rigby, R., Hayward, B., Liberatore, R. a., Bieniasz, P.D., Towers, G.J., Moita, L.F., Crow, Y.J., Bonthron, D.T., and Reis e Sousa, C. (2013). SAMHD1-dependent retroviral control and escape in mice. *EMBO J.* *32*, 2454–2462.
- Reichard, P. (1988). Interactions between deoxyribonucleotide and DNA synthesis. *Annu. Rev. Biochem.* *57*, 349–374.
- Rice, G.I., Bond, J., Asipu, A., Brunette, R.L., Manfield, I.W., Carr, I.M., Fuller, J.C., Jackson, R.M., Lamb, T., Briggs, T. a, Ali, M., Gornall, H., Couthard, L.R., Aeby, A., Attard-Montalto, S.P., Bertini, E., Bodemer, C., Brockmann, K., Brueton, L. a, Corry, P.C., Desguerre, I., Fazzi, E., Cazorla, A.G., Gener, B., Hamel, B.C.J., Heiberg, A., Hunter, M., van der Knaap, M.S., Kumar, R., Lagae, L., Landrieu, P.G., Lourenco, C.M., Marom, D., McDermott, M.F., van der Merwe, W., Orcesi, S., Prendiville, J.S., Rasmussen, M., Shalev, S. a, Soler, D.M., Shinawi, M., Spiegel, R., Tan, T.Y., Vanderver, A., Wakeling, E.L., Wassmer, E., Whittaker, E., Lebon, P., Stetson, D.B., Bonthron, D.T., and Crow, Y.J. (2009). Mutations involved in Aicardi-Goutières syndrome implicate SAMHD1 as regulator of the innate immune response. *Nat. Genet.* *41*, 829–832.
- Riebesell, U., Zondervan, I., Rost, B., Tortell, P.D., Zeebe, R.E., and Morel, F.M. (2000). Reduced calcification of marine plankton in response to increased atmospheric CO₂. *Nature* *407*, 364–367.
- Romanenko, S. a, Perelman, P.L., Trifonov, V. a, and Graphodatsky, a S. (2012). Chromosomal evolution in Rodentia. *Heredity (Edinb)*. *108*, 4–16.
- Rous, F. (1911). A Sarcoma of the Fowl Transmissible by an Agent Separable From the Tumor Cells.
- Rous, P. and Murphy, J.B. (1912). The nature of the filterable agent causing a sarcoma of the fowl. *J. Am. Med. Assoc.* *58*, 1938.

Rowe, H.M. and Trono, D. (2011). Dynamic control of endogenous retroviruses during development. *Virology* 411, 273–287.

Ryoo, J., Choi, J., Oh, C., Kim, S., Seo, M., Kim, S.-Y., Seo, D., Kim, J., White, T.E., Brandariz-Nuñez, A., Diaz-Griffero, F., Yun, C.-H., Hollenbaugh, J. a, Kim, B., Baek, D., and Ahn, K. (2014). The ribonuclease activity of SAMHD1 is required for HIV-1 restriction. *Nat. Med.* 20, 1–16.

Saavedra, C., Felber, B., and Izaurralde, E. (1997). The simian retrovirus-1 constitutive transport element, unlike the HIV-1 RRE, uses factors required for cellular mRNA export. *Curr. Biol.* 7, 619–628.

Sabin, A.B. (1938). Progression of Different Nasally Instilled Viruses along Different Nervous Pathways in the Same Host. *Exp. Biol. Med.* 38 , 270–275.

San Filippo, J., Sung, P., and Klein, H. (2008). Mechanism of eukaryotic homologous recombination. *Annu. Rev. Biochem.* 77, 229–257.

Sauerbrey, A., Stammer, G., Zintl, F., and Volm, M. (1998). Expression of the retinoblastoma tumor suppressor gene (RB-1) in acute leukemia. *Leuk. Lymphoma* 28, 275–283.

Sauter, D., Schindler, M., Specht, A., Landford, W.N., Kim, K., Votteler, J., Schubert, U., Bibollet-ruche, F., Keele, B.F., Takehisa, J., Ogando, Y., Ochsenbauer, C., John, C., Ayoub, A., Peeters, M., Learn, G.H., Shaw, G., Paul, M., Bieniasz, P., Hahn, B.H., Hatzioannou, T., and Kirchhoff, F. (2009). The evolution of pandemic and non-pandemic HIV-1 strains has been driven by Tetherin antagonism. *Cell Host Microbe* 6, 409–421.

Sayah, D.M., Sokolskaja, E., Berthou, L., and Luban, J. (2004). Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. *Nature* 430, 569–573.

Scarlata, S. and Carter, C. (2003). Role of HIV-1 Gag domains in viral assembly. *Biochim. Biophys. Acta - Biomembr.* 1614, 62–72.

Schärer, O.D. (2013). Nucleotide Excision Repair in Eukaryotes. *Cold Spring Harb. Perspect. Biol.* 5 .

Schlecht-Louf, G., Renard, M., Mangeney, M., Letzelter, C., Richaud, A., Ducos, B., Bouallaga, I., and Heidmann, T. (2010). Retroviral infection in vivo requires an immune escape virulence factor encrypted in the envelope protein of oncoretroviruses. *Proc. Natl. Acad. Sci. U. S. A.* 107, 3782–3787.

Schmidt, S., Schenkova, K., Adam, T., Erikson, E., Lehmann-koch, J., Sertel, S., Verhasselt, B., Fackler, O.T., Lasitschka, F., and Keppler, O.T. (2015). SAMHD1's protein expression profile in humans. *J. Leukoc. Biol.* 97, 1–10.

Schroeder, D.C., Oke, J., Malin, G., and Wilson, W.H. (2002). Coccolithovirus (Phycodnaviridae): Characterisation of a new large dsDNA algal virus that infects *Emiliania huxleyi*. *Arch. Virol.* 147, 1685–1698.

Schubert, U., Bour, S., Ferrer-Montiel, a V, Montal, M., Maldarell, F., and Strebel, K. (1996). The two biological activities of human immunodeficiency virus type 1 Vpu protein involve two separable structural domains. *J. Virol.* 70, 809–819.

- Segref, A., Sharma, K., Doye, V., Hellwig, A., Huber, J., Lührmann, R., and Hurt, E. (1997). Mex67p, a novel factor for nuclear mRNA export. Binds to both poly(A)⁺ RNA and nuclear pores. *EMBO J.* 16, 3256–3271.
- Selig, L., Pages, J.C., Tanchou, V., Prévéral, S., Berlioz-Torrent, C., Liu, L.X., Erdtmann, L., Darlix, J., Benarous, R., and Benichou, S. (1999). Interaction with the p6 domain of the gag precursor mediates incorporation into virions of Vpr and Vpx proteins from primate lentiviruses. *J. Virol.* 73, 592–600.
- Sharova, N., Wu, Y., Zhu, X., Stranska, R., Kaushik, R., Sharkey, M., and Stevenson, M. (2008). Primate lentiviral Vpx commandeers DDB1 to counteract a macrophage restriction. *PLoS Pathog.* 4.
- Sharp, P. (1983). Conversion of RNA to DNA in Mammals -ALU-Like Elements and Pseudogenes. *Nature* 301, 471–472.
- Sharp, P. and Hahn, B. (2011). Origins of HIV and the AIDS epidemic. *Cold Spring Harb Perspect Med.*
- Sheehy, A.M., Gaddis, N.C., Choi, J.D., and Malim, M.H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 418, 646–650.
- De Silva, S., Wang, F., Hake, T.S., Porcu, P., Wong, H.K., and Wu, L. (2013a). Downregulation of SAMHD1 Expression Correlates with Promoter DNA Methylation in Sézary Syndrome Patients. *J. Invest. Dermatol.* 134, 562–565.
- De Silva, S., Hoy, H., Hake, T.S., Wong, H.K., Porcu, P., and Wu, L. (2013b). Promoter methylation regulates SAMHD1 gene expression in human CD4⁺ T cells. *J. Biol. Chem.* 288, 9284–9292.
- Simon, F., Matheron, S., Tamalet, C., Loussertajaka, I., Bartczak, A., Pepin, J.M., Dhiver, C., Gamba, E., Elbim, C., A, G.J., Saimot, A.G., and Brunvezinet, F. (1993). Cellular and Plasma Viral Load in Patients Infected With HIV-2. *AIDS* 7, 1411–1417.
- Skubis-Zegadło, J., Stachurska, A., and Małecki, M. (2013). Vectorology of adeno-associated viruses (AAV). *Med. Wieku Rozwoj.* 17, 202–206.
- Sobhian, B., Laguette, N., Yatim, A., Nakamura, M., Levy, Y., Kiernan, R., and Benkirane, M. (2010). HIV-1 Tat Assembles a Multifunctional Transcription Elongation Complex and Stably Associates with the 7SK snRNP. *Mol. Cell* 38, 439–451.
- Solis, M., Nakhaei, P., Jalalirad, M., Lacoste, J., Douville, R., Arguello, M., Zhao, T., Laughrea, M., Wainberg, M. a, and Hiscott, J. (2011). RIG-I-mediated antiviral signaling is inhibited in HIV-1 infection by a protease-mediated sequestration of RIG-I. *J. Virol.* 85, 1224–1236.
- Springer, M.S. and Murphy, W.J. (2007). Mammalian evolution and biomedicine: New views from phylogeny. *Biol. Rev.* 82, 375–392.
- Srivastava, S., Swanson, S.K., Manel, N., Florens, L., Washburn, M.P., and Skowronski, J. (2008). Lentiviral Vpx accessory factor targets VprBP/DCAF1 substrate adaptor for cullin 4 E3 ubiquitin ligase to enable macrophage infection. *PLoS Pathog.* 4.
- St Gelais, C., de Silva, S., Amie, S.M., Coleman, C.M., Hoy, H., Hollenbaugh, J. a, Kim, B., and Wu, L. (2012). SAMHD1 restricts HIV-1 infection in dendritic cells (DCs) by dNTP depletion, but its expression in DCs and primary CD4⁺ T-lymphocytes cannot be upregulated by interferons. *Retrovirology* 9, 105.

- St Gelais, C., de Silva, S., Hach, J.C., White, T.E., Diaz-Griffero, F., Yount, J.S., and Wu, L. (2014). Identification of cellular proteins interacting with the retroviral restriction factor SAMHD1. *J. Virol.* **88**, 5834–5844.
- Stanley, T.L. and Grinspoon, S.K. (2012). Body composition and metabolic changes in HIV-infected patients. *J. Infect. Dis.* **205**, 383–390.
- Stopak, K., De Noronha, C., Yonemoto, W., and Greene, W.C. (2003). HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. *Mol. Cell* **12**, 591–601.
- Stremlau, M., Owens, C.M., Perron, M.J., Kiessling, M., Autissier, P., and Sodroski, J. (2004). The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. *Nature* **427**, 848–853.
- Stremlau, M., Perron, M., Lee, M., Li, Y., Song, B., Javanbakht, H., Diaz-Griffero, F., Anderson, D.J., Sundquist, W.I., and Sodroski, J. (2006). Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 5514–5519.
- Sundquist, W.I. and Kra, H. (2012). HIV-1 Assembly, Budding, and Maturation. 1–24.
- Suzuki, H., Maruyama, R., Yamamoto, E., and Kai, M. (2012). DNA methylation and microRNA dysregulation in cancer. *Mol. Oncol.* **6**, 567–578.
- Sze, A., Belgnaoui, S.M., Olganier, D., Lin, R., Hiscott, J., and Van Grevenynghe, J. (2013). Host restriction factor SAMHD1 limits human T cell leukemia virus type 1 infection of monocytes via STING-mediated apoptosis. *Cell Host Microbe* **14**, 422–434.
- Tabler, C.O., Lucera, M.B., Haqqani, A. a, McDonald, D.J., Migueles, S. a, Connors, M., and Tilton, J.C. (2014). CD4+ memory stem cells are infected by HIV-1 in a manner regulated in part by SAMHD1 expression. *J. Virol.* **88**, 4976–4986.
- Tatum, E.L. (1966). Molecular biology, nucleic acids, and the future of medicine. *Perspect. Biol. Med.* **10**, 19–32.
- Temin, H.M. and Mizutani, S. (1970). RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* **226**, 1211–1213.
- The Journal of Gene Medicine Clinical Trial site.
- Thiele, H., Du Moulin, M., Barczyk, K., George, C., Schwindt, W., Nürnberg, G., Frosch, M., Kurlemann, G., Roth, J., Nürnberg, P., and Rutsch, F. (2010). Cerebral arterial stenoses and stroke: novel features of Aicardi-Goutières syndrome caused by the Arg164X mutation in SAMHD1 are associated with altered cytokine expression. *Hum. Mutat.* **31**, 1836–1850.
- Thomas, F. and Kutay, U. (2003). Biogenesis and nuclear export of ribosomal subunits in higher eukaryotes depend on the CRM1 export pathway. *J. Cell Sci.* **116**, 2409–2419.
- Tiscornia, G., Singer, O., and Verma, I.M. (2006). Production and purification of lentiviral vectors. *Nat. Protoc.* **1**, 241–245.

- Tokarev, A., Suarez, M., Kwan, W., Fitzpatrick, K., Singh, R., and Guatelli, J. (2013). Stimulation of NF- κ B activity by the HIV restriction factor BST2. *J. Virol.* 87, 2046–2057.
- Trilling, D.M. and Axelrod, D. (1970). Encapsidation of free host DNA by simian virus 40: a simian virus 40 pseudovirus. *Science* 168, 268–271.
- Trono, D. (2000). Lentiviral vectors: turning a deadly foe into a therapeutic agent. *Gene Ther.* 7, 20–23.
- Tüngler, V., Staroske, W., Kind, B., Dobrick, M., Kretschmer, S., Schmidt, F., Krug, C., Lorenz, M., Chara, O., Schwille, P., and Lee-Kirsch, M.A. (2013). Single-stranded nucleic acids promote SAMHD1 complex formation. *J. Mol. Med.* 91, 759–770.
- Ugolini, G. (1995). Specificity of rabies virus as a transneuronal tracer of motor networks: Transfer from hypoglossal motoneurons to connected second-order and higher order central nervous system cell groups. *J. Comp. Neurol.* 356, 457–480.
- Ugolini, G., Kuypers, H.G., and Simmons, a (1987). Retrograde transneuronal transfer of herpes simplex virus type 1 (HSV 1) from motoneurons. *Brain Res.* 422, 242–256.
- UNAIDS (2014). Ending the AIDS epidemic. Fact Sheet.
- VanDamme, N., Goff, D., Katsura, C., Jorgenson, R., Mitchell, R., Johnson, M., Stephens, E., and Guatelli, J. (2008). The interferon-induced protein BST-2/CD317 restricts release of virions from infected cells and is down-regulated from the cell surface by HIV-1 Vpu. *Cell Host Microbe.*
- Verbrugge, I., Johnstone, R.W., and Smyth, M.J. (2010). SnapShot: Extrinsic apoptosis pathways. *Cell* 143, 1192–1192.e2.
- Verheggen, C. and Bertrand, E. (2012). CRM1 plays a nuclear role in transporting snoRNPs to nucleoli in higher eukaryotes. *Nucleus* 3, 19–24.
- Vogel, F. (1979). Genetics of Retinoblastoma. *Hum. Genet.* 54, 1–54.
- Vulin, A., Barthélémy, I., Goyenvallé, A., Thibaud, J.-L., Beley, C., Griffith, G., Benchaouir, R., le Hir, M., Unterfinger, Y., Lorain, S., Dreyfus, P., Voit, T., Carlier, P., Blot, S., and Garcia, L. (2012). Muscle Function Recovery in Golden Retriever Muscular Dystrophy After AAV1-U7 Exon Skipping. *Mol. Ther.* 20, 2120–2133.
- Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S., and Alizon, M. (1985). Nucleotide sequence of the AIDS virus, LAV. *Cell* 40, 9–17.
- Wang, J., Lu, F., Shen, X.-Y., Wu, Y., and Zhao, L. (2014a). SAMHD1 is down regulated in lung cancer by methylation and inhibits tumor cell proliferation. *Biochem. Biophys. Res. Commun.* 455, 229–233.
- Wang, S., Han, X., Wang, J., Yao, J., and Shi, Y. (2014b). Antitumor Effects of a Novel Chromosome Region Maintenance 1 (CRM1) Inhibitor on Non-Small Cell Lung Cancer Cells In Vitro and in Mouse Tumor Xenografts. *PLoS One* 9, e89848.
- Weber, G.J., Sepúlveda, M.S., Peterson, S.M., Lewis, S.S., and Freeman, J.L. (2013). Transcriptome alterations following developmental atrazine exposure in zebrafish are associated with disruption of

neuroendocrine and reproductive system function, cell cycle, and carcinogenesis. *Toxicol. Sci.* 132, 458–466.

Weber, M., Rabinowitz, J., Provost, N., Conrath, H., Folliot, S., Briot, D., Chérel, Y., Chenuaud, P., Samulski, J., Moullier, P., and Rolling, F. (2003). Recombinant adeno-associated virus serotype 4 mediates unique and exclusive long-term transduction of retinal pigmented epithelium in rat, dog, and nonhuman primate after subretinal delivery. *Mol. Ther.* 7, 774–781.

Welbourn, S., Dutta, S.M., Semmes, O.J., and Strebel, K. (2013). Restriction of Virus Infection but Not Catalytic dNTPase Activity Is Regulated by Phosphorylation of SAMHD1. *J. Virol.* 87, 11516–11524.

White, T.E., Brandariz-Núñez, A., Carlos Valle-Casuso, J., Amie, S., Nguyen, L., Kim, B., Brojatsch, J., and Diaz-Griffero, F. (2013a). Contribution of SAM and HD domains to retroviral restriction mediated by human SAMHD1. *Virology* 436, 81–90.

White, T.E., Brandariz-Núñez, A., Valle-Casuso, J.C., Amie, S., Nguyen, L.A., Kim, B., Tuzova, M., and Diaz-Griffero, F. (2013b). The retroviral restriction ability of SAMHD1, but not its deoxynucleotide triphosphohydrolase activity, is regulated by phosphorylation. *Cell Host Microbe* 13, 441–451.

Wickersham, I.R., Lyon, D.C., Barnard, R.J.O., Mori, T., Conzelmann, K., Young, J. a T., Callaway, E.M., Manuscript, A., and Neurons, G.T. (2009). Monosynaptic Restriction of Transsynaptic Tracing from Single, Genetically Targeted Neurons. *Neuron.* 53, 639–647.

Widłak, P. (2000). The DFF40/CAD endonuclease and its role in apoptosis. *Acta Biochim. Pol.* 47, 1037–1044.

Wiegand, H.L., Coburn, G. a, Zeng, Y., Kang, Y., Bogerd, H.P., and Cullen, B.R. (2002). Formation of Tap / NXT1 Heterodimers Activates Tap-Dependent Nuclear mRNA Export by Enhancing Recruitment to Nuclear Pore Complexes Formation of Tap / NXT1 Heterodimers Activates Tap-Dependent Nuclear mRNA Export by Enhancing Recruitment to Nuclear Pore C. *Society* 22, 245–256.

Wilson, W.H., Tarran, G. a., Schroeder, D., Cox, M.J., Oke, J., and Malin, G. (2002). Isolation of viruses responsible for the demise of an *Emiliana huxleyi* bloom in the English Channel. 369–377.

Wirth, T., Parker, N., and Ylä-Herttuala, S. (2013). History of gene therapy. *Gene* 525, 162–169.

Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., and Micklem, G. (1995). Identification of the breast cancer susceptibility gene BRCA2. *Nature* 378, 789–792.

Wu, X., Anderson, J.L., Campbell, E.M., Joseph, A.M., and Hope, T.J. (2006). Proteasome inhibitors uncouple rhesus TRIM5alpha restriction of HIV-1 reverse transcription and infection. *Proc. Natl. Acad. Sci. U. S. A.* 103, 7465–7470.

Xin, B., Jones, S., Puffenberger, E.G., Hinze, C., Bright, A., Tan, H., Zhou, A., Wu, G., Vargus-Adams, J., Agamanolis, D., and Wang, H. (2011). Homozygous mutation in SAMHD1 gene causes cerebral vasculopathy and early onset stroke. *Proc. Natl. Acad. Sci. U. S. A.* 108, 5372–5377.

Yan, J., Kaur, S., DeLucia, M., Hao, C., Mehrens, J., Wang, C., Golczak, M., Palczewski, K., Gronenborn, A.M., Ahn, J., and Skowronsk, S. (2013). Tetramerization of SAMHD1 is required for biological activity and inhibition of HIV infection. *J. Biol. Chem.* 288, 10406–10417.

- Yan, N., Regalado-Magdos, A.D., Stiggelbout, B., Lee-Kirsch, M.A., and Lieberman, J. (2010). The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1. *Nat. Immunol.* **11**, 1005–1013.
- Yang, J., Bill, M. a., Young, G.S., La Perle, K., Landesman, Y., Shacham, S., Kauffman, M., Senapedis, W., Kashyap, T., Saint-Martin, J.R., Kendra, K., and Lesinski, G.B. (2014a). Novel small molecule XPO1/CRM1 inhibitors induce nuclear accumulation of TP53, phosphorylated MAPK and apoptosis in human melanoma cells. *PLoS One* **9**.
- Yang, S., Shan, T., Zhou, Y., Jiang, Y., Tong, W., Liu, F., Wen, F., Zhang, Q., and Tong, G. (2014b). Molecular cloning and characterizations of porcine SAMHD1 and its roles in replication of highly pathogenic porcine reproductive and respiratory syndrome virus. *Dev. Comp. Immunol.* **47**, 234–246.
- Yu, J. and Zhang, L. (2005). The transcriptional targets of p53 in apoptosis control. *Biochem. Biophys. Res. Commun.* **331**, 851–858.
- Yu, C., Shen, K., Lin, M., Chen, P., Lin, C., Chang, G.D., and Chen, H. (2002). GCMa regulates the syncytin-mediated trophoblastic fusion. *J. Biol. Chem.* **277**, 50062–50068.
- Yu, H., Usmani, S.M., Borch, A., Krämer, J., Stürzel, C.M., Khalid, M., Li, X., Krnavek, D., van der Ende, M.E., Osterhaus, A.D., Gruters, R. a, and Kirchhoff, F. (2013). The efficiency of Vpx-mediated SAMHD1 antagonism does not correlate with the potency of viral control in HIV-2-infected individuals. *Retrovirology* **10**, 27.
- Yu, X., Yu, Y., Liu, B., Luo, K., Kong, W., Mao, P., and Yu, X.-F. (2003). Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* **302**, 1056–1060.
- Zane, L. and Jeang, K.-T. (2014). HTLV-1 and leukemogenesis: virus-cell interactions in the development of adult T-cell leukemia. *Recent Results Cancer Res.* **193**, 191–210.
- Zemanick, M.C., Strick, P.L., and Dix, R.D. (1991). Direction of transneuronal transport of herpes simplex virus 1 in the primate motor system is strain-dependent. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 8048–8051.
- Zhang, H., Yang, B., Pomerantz, R.J., Zhang, C., Arunachalam, S.C., and Gao, L. (2003). The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **424**, 94–98.
- Zhang, W., Du, J., Evans, S.L., Yu, Y., and Yu, X.-F. (2011). T-cell differentiation factor CBF- β regulates HIV-1 Vif-mediated evasion of host restriction. *Nature* **481**, 376–379.
- Zhao, K., Du, J., Han, X., Goodier, J.L., Li, P., Zhou, X., Wei, W., Evans, S.L., Li, L., Zhang, W., Cheung, L.E., Wang, G., Kazazian, H.H., and Yu, X.F. (2013). Modulation of LINE-1 and Alu/SVA Retrotransposition by Aicardi-Goutières Syndrome-Related SAMHD1. *Cell Rep.* **4**, 1108–1115.
- Zheng, N.N., Kiviat, N.B., Sow, P.S., Hawes, S.E., Wilson, a, Critchlow, C.W., Gottlieb, G.S., Musey, L., and Mcelrath, M.J. (2004a). Comparison of Human Immunodeficiency Virus (HIV) -Specific T-Cell Responses in HIV-1- and HIV-2-Infected Individuals in Senegal. **78**, 13934–13942.
- Zheng, Y., Irwin, D., Kurosu, T., Sata, T., Peterlin, B.M., and Tokunaga, K. (2004b). Human APOBEC3F Is Another Host Factor That Blocks Human Immunodeficiency Virus Type 1 Replication. *J. Virol.* **78**, 6073–6076.

Zhou, X., Evans, S.L., Han, X., Liu, Y., and Yu, X.F. (2012). Characterization of the interaction of full-length HIV-1 Vif protein with its key regulator CBF β and CRL5 E3 ubiquitin ligase components. *PLoS One* 7, 1–10.

Zhu, C., Gao, W., Zhao, K., Qin, X., Zhang, Y., Peng, X., Zhang, L., Dong, Y., Zhang, W., Li, P., Wei, W., Gong, Y., and Yu, X.-F. (2013). Structural insight into dGTP-dependent activation of tetrameric SAMHD1 deoxynucleoside triphosphate triphosphohydrolase. *Nat. Commun.* 4, 2722.