



**HAL**  
open science

# Epigénétique et méthylation de l'ADN : étude des mécanismes d'interaction du domaine SRA de UHRF1 avec l'ADN hémi-méthylé

Vanille Greiner

► **To cite this version:**

Vanille Greiner. Epigénétique et méthylation de l'ADN : étude des mécanismes d'interaction du domaine SRA de UHRF1 avec l'ADN hémi-méthylé. Génomique, Transcriptomique et Protéomique [q-bio.GN]. Université de Strasbourg, 2012. Français. NNT : 2012STRAJ107 . tel-01124072

**HAL Id: tel-01124072**

**<https://theses.hal.science/tel-01124072v1>**

Submitted on 6 Mar 2015

**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.

*ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE*

UMR CNRS 7213

**THÈSE** présentée par :

**Vanille GREINER**

soutenue le : 13 décembre 2012

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Sciences du Vivant

**Epigénétique et méthylation de l'ADN :  
Etude des mécanismes d'interaction du  
domaine SRA de UHRF1 avec l'ADN  
hémi-méthylé**

**THÈSE dirigée par :**

**M. MELY Yves**

**M. BRONNER Christian**

Professeur, Université de Strasbourg

Chargé de Recherches, Université de Strasbourg

**RAPPORTEURS :**

**M. DANTE Robert**

**M. MARGEAT Emmanuel**

Directeur de Recherches, Centre Léon Bérard

Chargé de Recherches, Centre de Biochimie Structurale

---

**AUTRES MEMBRES DU JURY :**

**M. CARTRON Pierre-François**

**M. CAVARELLI Jean**

Chargé de Recherches, Centre de Recherche en Cancérologie

Professeur, Université de Strasbourg



*Je tiens à témoigner toute ma reconnaissance au Docteur Robert Dante et au Docteur Emmanuel Margeat pour avoir accepté d'être rapporteurs de ma thèse et d'avoir consacré du temps à la lecture, l'analyse, et la critique de ce travail. Je voudrais également remercier le Professeur Jean Cavarelli et le Docteur Pierre-François Cartron pour m'avoir fait l'honneur de participer à ce jury de thèse.*

*Je remercie sincèrement mon directeur de thèse, le Professeur Yves Mély pour m'avoir accueillie dans son laboratoire et pour m'avoir fait confiance tout au long de mon travail. Mille mercis pour votre disponibilité, votre soutien, et vos conseils scientifiques qui ont été riches d'enseignement.*

*J'adresse également mes remerciements à mon co-directeur de thèse, le Docteur Christian Bronner, pour ses conseils et son expertise scientifique.*

*Je remercie Sanofi Pasteur d'avoir permis le financement de ma thèse, et plus particulièrement Frédéric Ronzon et Catherine Manin pour nos échanges et notre sympathique collaboration.*

*Je remercie Hugues de Rocquigny qui m'a permis de découvrir le monde de la recherche. Merci pour vos conseils, votre soutien et nos échanges.*

*Merci à Catherine Birck pour ses conseils, et pour la production et la purification de la protéine UHRF1.*

*Merci à Nicolas Lévy pour sa gentillesse et sa disponibilité, et pour la purification de la protéine UHRF1 après marquage.*

*Merci à nos supers secrétaires, Ingrid Barthel et Marlyse Wernert, toujours disponibles et toujours souriantes.*

*Merci à Ludovic Four pour sa disponibilité et ses sauvetages informatiques.*

*Merci à Andrey pour sa gentillesse, sa disponibilité et son enthousiasme.*

*Merci à Joëlle, "ma grande sœur du labo", qui m'a beaucoup appris.*

*Merci à Julien, pour son aide, pour tous ses conseils scientifiques et la richesse de nos discussions. J'ai beaucoup appris grâce à toi.*

*Merci à Noémie, "ma petite sœur du labo" avec qui j'ai beaucoup partagé. Je n'oublierai pas nos pauses café, nos discussions, nos fous rires, nos encouragements mutuels ("Eye of the Tiger"). Je te souhaite pleins de bonheur et d'épanouissement.*

*Merci à Manu pour sa bonne humeur, sa gentillesse, sa fraîcheur, son écharpe, et son esprit "Robin des bois" dans lequel je me retrouve tant. En grand merci aussi à Eléonore, d'être entière et sincère, merci pour tes coups de gueule et ta bonne humeur (si, si), merci pour tes conseils. Merci à vous deux pour votre soutien, nos discussions, nos débats, et pour tous ces agréables moments passés ensemble.*

*Un immense merci à Youri, qui est devenu un ami et un confident. Merci pour ton soutien sans faille, ta gentillesse, tes blagues, tes encouragements, ... Merci pour tout.*

*Merci à Pascal pour ses conseils, son enseignement,... et ses filouteries.*

*Merci à Nicolas pour sa gentillesse et sa disponibilité. Je n'oublierai pas nos enquêtes "cuvette" et nos sessions coup de gueule. Je te souhaite bonne continuation dans ce monde "souverain" ; )*

*Merci à Avisek pour son sourire, sa bonne humeur sans faille et sa gentillesse. Merci également à mes autres collègues du "mini zoo" qui ont participé à la bonne ambiance de cette pièce, Hussein, et Rémy pour son soutien technique.*

*Merci à Ludovic Richert pour sa disponibilité et sa "magic touch".*

*Merci à Annie Wund et Nicole Glasser de m'avoir initiée à l'enseignement. Merci pour vos conseils et votre gentillesse.*

*Je tiens à remercier chaleureusement tous les autres membres de mon équipe Halina, Frédéric ("easy"), Christian Boudier, Guy Duportail, Frédéric, Andreas, ... et les collègues du zoo Zeinab, Hala, Armelle, Kamal, Ievgen, Sarwat,...*

*Je tiens à remercier tout particulièrement mes amis qui ont toujours été présents, et qui m'ont encouragée et soutenue tout au long de ma thèse. Je sais à quel point j'ai de la chance d'être si bien entourée et d'avoir des amis tels que vous. Merci à la super équipe Briçou, Nico, Théo, Léli, Bléonard,... et merci à mes autres soutiens Diane, Sophie,...*

*Merci à ma petite Nona qui m'a toujours soutenue. Merci à ce super petit bout de femme de caractère sans qui j'aurais eu du mal à passer bien des épreuves.*

*Merci à ma "dream team", Ludo (le petit rayon de soleil), Marc, ma p'tite Aude, Aurore, Ced. Merci de me soutenir et de m'encourager. Merci pour tous ces inoubliables moments passés ensemble et pour tous les moments futurs...*

*Merci à toute ma famille qui m'a toujours soutenue et encouragée. Merci à mes supers tatas et tontons, Chantal, Folke, Paulo, Lison ; mes supers cousins, Nils, Matt, Jérémie, Olivier, et toutes leur petites familles Laurence, Céline, Nana, Jojo, Gustave, Elise, Juliette ; merci à ma cousine "lait miel", Frantz, Fanny, Bernard... Merci pour tous ces bons moments passés en famille qui m'ont toujours si bien ressourcés.*

*Une énorme pensée pour tante France et ma petite mamie qui auraient été si fières et qui auraient aimé partager ce moment avec nous.*

*Merci à mon super beau-papa, pour ses "petits verres", sa bonne humeur, ses bons plats et ses débats.*

*Et enfin et surtout, merci à mon père et à ma mère pour leur amour et leur soutien inconditionnel...*

## LISTE DES ABREVIATIONS

<b>2-Ap</b>	2-aminopurine
<b>5cac</b>	5-carboxylcytosine
<b>5fC</b>	5-formylcytosine
<b>5hmC</b>	5-hydroxyméthylcytosine
<b>5hmU</b>	5-hydroxyméthyluracile
<b>5mC</b>	5-méthylcytosine
<b>AdoHcy</b>	S-adenosylhomocystéine (= SAH)
<b>AdoMet</b>	S-adénosylméthionine (=SAM)
<b>AID</b>	"activation-induced cytidine deaminase"
<b>APOBEC</b>	"apolipoprotein B mRNA editing enzyme, catalytic polypeptide"
<b>ARNm</b>	ARN messenger
<b>ARNnc</b>	ARN non-codant
<b>ARNpi</b>	ARN interagissant avec Piwi
<b>ATM</b>	"ataxia telangectasia mutated"
<b>ATR</b>	"ATM and Rad3-related"
<b>ATRX</b>	"alpha thalassaemia/mental retardation X-linked"
<b>BER</b>	"base excision repair", réparation par excision de base
<b>cAMP</b>	adénosine monophosphate cyclique
<b>CAPs</b>	"Chromatin Architectural Proteins", protéines d'architecture de la chromatine
<b>Cdk</b>	kinases cycline-dépendantes
<b>CHAPS</b>	"3-[(3-cholamidopropyl)diméthylammonio]-1-propanesulfonic acid"
<b>CREBBP</b>	"cAMP-response element binding protein"
<b>DDT</b>	"DNA binding homeobox and Different Transcription factors"
<b>DGCR8</b>	"Di George Critical Region 8"
<b>DMF</b>	diméthylformamide
<b>DNMT</b>	méthyltransférase de l'ADN
<b>FDA</b>	"Food and Drug Administration"
<b>FRET</b>	"Fluorescence Resonance Energy Transfer", transfert d'énergie non-radiatif
<b>gDMRs</b>	"germline differentially methylated regions"

<b>HAT</b>	acétyltransférase des histone s
<b>HAT-MBD</b>	"histone acetyltransferases methyl-binding domain"
<b>HAUSP</b>	"herpes virus-associated ubiquitin specific protease" (= Usp7)
<b>HDAC</b>	déacétylase des histones
<b>HDM</b>	déméthylase des histones
<b>HMT</b>	méthyltransférase des histones
<b>HMT-MBD</b>	"histone methyltransferases methyl-binding domain"
<b>HP1</b>	protéine de l'hétérochromatine 1
<b>IAPs</b>	"intracisternal A-type particles"
<b>ICR</b>	"imprinting control region", régions de contrôle d'empreinte ICR
<b>IDF</b>	"immunodeficiency, centromeric instability and facial anomalies"
<b>IPTG</b>	isopropyl $\beta$ -D-1-thiogalactopyranoside
<b>Kd</b>	constante de dissociation
<b>LB</b>	Lysogeny broth
<b>LincARN</b>	large ARN non-codant intergénique
<b>lncARN</b>	long ARN non-codant
<b>MBD</b>	"methyl-CpG binding domain"
<b>MBD4</b>	"methyl-CpG-binding domain protein 4"
<b>MBP</b>	"methyl-binding protein"
<b>miARN</b>	microARN
<b>NATA</b>	N-acétyle-L-tryptophanamide
<b>NCoR</b>	"nuclear receptor corepressor"
<b>NoRC</b>	"nucleolar remodeling complex"
<b>PARP</b>	poly-ADPribose polymérase
<b>PCNA</b>	"proliferating cell nuclear antigen"
<b>PEG</b>	polyéthylène glycol
<b>PEU</b>	photoélectron unique
<b>PHD</b>	"Plant Homeo Domain"
<b>PKA</b>	protéine kinase A
<b>PKMT</b>	protéine lysine métyltransférase
<b>PML</b>	"promyelocytic leukemia"
<b>PMSF</b>	fluorure de phénylméthylsulfonyl
<b>pRb</b>	protéine du rétinoblastome
<b>PRMT</b>	protéine arginine métyltransférase



<b>RdDM</b>	"RNA-directed DNA methylation", méthylation de l'ADN dirigée par l'ARN
<b>RING</b>	"Really Interesting New Gene domain"
<b>RISC</b>	"RNA-induced silencing complex"
<b>RNP</b>	"RNA protein complexes"
<b>rTEV</b>	"recombinant tobacco etch virus"
<b>SAH</b>	S-adenosylhomocystéine (= AdoHcy)
<b>SAM</b>	S-adenosylméthionine (= AdoMet)
<b>SLA</b>	sclérose latérale amyotrophique
<b>SMUG1</b>	"single-strand-selective monofunctional uracil-DNA glycosylase 1"
<b>SRA</b>	"SET and RING Associated domain"
<b>SUMO</b>	"small ubiquitin-like modifier"
<b>TCEP</b>	Tris(2-carboxyéthyl)phosphine
<b>TCR</b>	récepteur des cellules T
<b>TDG</b>	"thymine DNA glycosylase", glycosylase de la thymine de l'ADN
<b>T-DMR</b>	"tissue specific differentially methylated region"
<b>Tet</b>	"ten-eleven-translocation"
<b>TGS</b>	"transcriptional gene silencing"
<b>Trp</b>	tryptophane
<b>TTD</b>	"cryptic Tandem Tudor Domain"
<b>UBL</b>	"Ubiquitin-like domain" (= NIRF-N)
<b>UHRF1</b>	"Ubiquitin-like plant Homeodomain and Ring Finger 1"
<b>Usp7</b>	"ubiquitin specific peptidase 7" (= HAUSP)
<b>UTR</b>	"untranslated region"
<b>Xi</b>	chromosome X inactivé
<b>XPO5</b>	Exportin 5

# TABLE DES MATIERES

<b>I- INTRODUCTION</b> .....	4
A. Les modifications épigénétiques .....	5
1. La chromatine, support de l'information épigénétique .....	5
1.1. Composition de la chromatine.....	5
1.2. Structure de la chromatine.....	6
1.3. Euchromatine et hétérochromatine.....	9
2. La méthylation de l'ADN.....	10
2.1. Méthylation par les méthyltransférases de l'ADN.....	10
2.2. Les sites de méthylation de l'ADN .....	15
2.3. Les modules de reconnaissance des CpG méthylés .....	18
2.4. La méthylation <i>de novo</i> .....	21
2.5. La méthylation de maintenance.....	26
2.6. La déméthylation de l'ADN.....	27
2.7. Rôles physiologiques de la méthylation de l'ADN .....	30
2.8. Dérégulation de la méthylation de l'ADN .....	33
3. Les modifications post-traductionnelles des histones .....	34
3.1. L'acétylation des histones.....	36
3.2. La méthylation des histones .....	37
3.3. Autres modifications post-traductionnelles des histones .....	38
3.4. Le code histone.....	39
3.5. Connexion entre la méthylation de l'ADN et les modifications des histones....	41
3.6. Dérégulation des modifications des histones .....	43
4. Les variants d'histones.....	44
5. Positionnement des nucléosomes .....	46
6. Les ARNs non-codant .....	47
B. Epigénétique et cancer .....	50
1. Hyperméthylation locale versus hypométhylation globale .....	51
2. Modifications des histones dans le cancer .....	53
3. Thérapie épigénétique .....	53
C. La protéine humaine hUHRF1 et la régulation épigénétique.....	55
1. La famille UHRF .....	56

2.	Architecture de la protéine hUHRF1 .....	57
2.1.	Le domaine UBL ("Ubiquitin-like domain") .....	57
2.2.	Le domaine TTD ("cryptic Tandem Tudor Domain") .....	58
2.3.	Le domaine PHD ("Plant Homeo Domain") .....	59
2.4.	Le domaine SRA ("Set and Ring Associated domain") .....	60
2.5.	Le domaine RING ("Really Interesting New Gene domain") .....	64
3.	Expression cellulaire de hUHRF1 .....	65
4.	Les différents rôles de la protéine hUHRF1 .....	66
4.1.	Auto-régulation de UHRF1 et régulation de la Dnmt1 .....	66
4.2.	hUHRF1 dans la régulation du cycle cellulaire .....	68
4.3.	hUHRF1 dans la régulation épigénétique .....	72
4.4.	hUHRF1 dans la cancérogenèse .....	74
5.	UHRF1 comme marqueur diagnostique et pronostique du cancer .....	75
6.	UHRF1 comme cible thérapeutique potentielle .....	76
<b>II-</b>	<b>OBJECTIFS</b> .....	<b>77</b>
<b>III-</b>	<b>MATERIELS ET METHODES</b> .....	<b>79</b>
A.	Matériels .....	79
1.	Domaine SRA (408-643) de la protéine UHRF1 .....	79
1.1.	Production et purification de la protéine .....	79
1.2.	Stockage de la protéine .....	80
1.3.	Marquage de la protéine au Cy3 .....	80
2.	Séquences oligonucléotidiques .....	84
2.1.	Séquences utilisées .....	84
2.2.	Mise en solution .....	85
2.3.	Hybridation des oligonucléotides .....	85
3.	Conditions expérimentales des mesures des interactions protéine/ADN .....	86
B.	Méthodes .....	86
1.	Spectroscopie d'absorption UV-visible .....	86
1.1.	Principe et appareillage .....	86
1.2.	Courbes de fusion des oligonucléotides .....	87
2.	Spectroscopie de fluorescence à l'état stationnaire .....	89
2.1.	Emission de fluorescence .....	89
2.1.1.	Principe et appareillage .....	89

2.1.2.	Fluorescence du tryptophane .....	90
2.1.3.	Fluorescence de la 2-aminopurine .....	92
2.1.4.	Mesure du rendement quantique de fluorescence .....	93
2.1.5.	Mesure de l'interaction de la protéine SRA aux acides nucléiques par transfert d'énergie non-radiatif (FRET) .....	94
2.2.	Anisotropie de fluorescence .....	98
2.2.1.	Principe et appareillage .....	98
2.2.2.	Mesure de la liaison des protéines SRA aux acides nucléiques par titrage direct .....	100
3.	Spectroscopie de fluorescence résolue en temps.....	101
3.1.	Principes de la technique du photoélectron unique et dispositif de mesure.....	102
3.2.	Mesure des durées de vie de fluorescence .....	105
3.3.	Mesure des déclins d'anisotropie de fluorescence .....	106
3.4.	Conditions expérimentales .....	108
<b>IV-</b>	<b>RESULTATS ET DISCUSSION</b> .....	<b>109</b>
A.	Publication 1 : "Site-specific characterization of the dynamics of UHRF1 SRA domain bound to 2-aminopurine labeled hemi-methylated oligonucleotides" .....	109
B.	Caractéristiques spectroscopiques du domaine SRA (408-643) .....	144
1.	Tests de conservation de la protéine .....	144
2.	Problèmes d'adsorption sur les parois des cuves de mesure .....	145
3.	Caractéristiques spectroscopiques de la protéine .....	147
C.	Liaison du domaine SRA (408-643) à l'ADN .....	149
D.	Liaison du domaine SRA (408-643) à l'ARN.....	152
E.	Effet de la 5-azacytidine sur la liaison SRA/ADN.....	154
<b>V-</b>	<b>CONCLUSION ET PERSPECTIVES</b> .....	<b>156</b>
<b>VI-</b>	<b>TRAVAUX ANNEXES</b> .....	<b>160</b>
A.	Publication 2 : "Characterization of the lipid and protein organization in HBsAg viral particles by steady-state and time-resolved fluorescence spectroscopy" .....	160
B.	Publication 3 : "The structure of HBsAg particles is not modified upon their adsorption on aluminium hydroxide gel" .....	161
C.	Publication 4 : "Characterization of the mechanisms of HIV-1 Vpr(52-96) internalization in cells" .....	162
<b>VII-</b>	<b>REFERENCES BIBLIOGRAPHIQUES</b> .....	<b>163</b>

## I- INTRODUCTION

A l'exception des cellules germinales, toutes les cellules d'un organisme possèdent le même code génétique et le même nombre de gènes. Pourtant, lorsqu'elles se différencient, les cellules vont acquérir des profils d'expression caractéristiques d'un tissu et exprimer les gènes qui sont seuls nécessaires à leurs fonctions. L'existence d'épigénomes distincts explique pourquoi le même génotype peut générer différents phénotypes. Ce sont les modifications épigénétiques, héritables lors des divisions cellulaires, qui régissent ainsi le destin d'une cellule, en régulant l'expression de ses gènes sans affecter la séquence de l'ADN (Berger et al., 2009). On distingue ainsi, la méthylation de l'ADN, les modifications post-traductionnelles des histones, les changements de variants d'histones, et les ARNs non-codant, qui vont entraîner des modifications dans l'architecture de la chromatine, support de l'information épigénétique. Ces mécanismes sont régulés par un grand nombre d'acteurs capables d'écrire, de lire ou d'effacer ces modifications et qui vont travailler en coopération pour inhiber ou activer l'expression de gènes spécifiques. L'épigénétique joue un rôle essentiel dans un grand nombre de mécanismes fondamentaux tels que le développement embryonnaire, la différenciation cellulaire et l'identité cellulaire (Smallwood and Kelsey, 2012). On découvre aujourd'hui que les altérations épigénétiques sont ainsi impliquées dans un grand nombre de pathologies telles que le cancer, les maladies neurologiques ou auto-immunes (Portela and Esteller, 2010). Les modifications épigénétiques étant réversibles, la compréhension des mécanismes régissant leur régulation ouvre la porte à de nouvelles stratégies thérapeutiques qui deviendront certainement dans le futur des outils indispensables pour combattre la maladie.

Parmi les acteurs de la régulation épigénétique, la protéine UHRF1 joue un rôle central dans ce processus puisqu'elle se situe à l'interface des modifications de l'ADN et des histones. Elle possède ainsi un domaine SRA (Set and Ring Associated) capable de reconnaître l'ADN hémi-méthylé (Arita et al., 2008; Avvakumov et al., 2008; Bostick et al., 2007; Frauer et al., 2011; Hashimoto et al., 2008; Qian et al., 2008; Rottach et al., 2010 ; Sharif et al., 2007; Unoki et al., 2004) et les domaines TTD et PHD susceptibles de lire les modifications des histones (Arita et al., 2012; Nady et al., 2011; Xie et al., 2012). En recrutant divers acteurs de la régulation épigénétique au niveau de régions spécifiques de la chromatine, la protéine UHRF1 est ainsi capable de maintenir et de transmettre l'information épigénétique

(Bronner et al., 2010; Hashimoto et al., 2009; Hashimoto et al., 2010). L'étude de cette protéine apporte donc un nouveau niveau de compréhension dans les mécanismes épigénétiques. D'autre part, la protéine UHRF1 étant impliquée dans le développement de certaines tumeurs, elle constituerait une nouvelle cible thérapeutique potentielle dans le traitement des cancers.

## A. Les modifications épigénétiques

Par définition, les modifications épigénétiques constituent des changements dans l'expression des gènes causés par des mécanismes qui n'affectent pas la séquence de l'ADN et qui sont transmissibles à la descendance (Berger et al., 2009). Elles peuvent être de plusieurs ordres, la méthylation de l'ADN, les modifications post-traductionnelles des histones, les variants d'histones, les complexes de remodelage de la chromatine ATP-dépendants et les ARNs non-codant, la chromatine constituant le support de ces modifications épigénétiques.

### 1. La chromatine, support de l'information épigénétique

Dans les cellules eucaryotes, l'ADN est présent dans le noyau sous forme d'une structure nucléoprotéique appelée chromatine. Outre sa fonction d'empaquetage, qui permet de condenser l'ADN de 2 m de long dans un noyau de 6  $\mu\text{m}$  de diamètre en moyenne, la chromatine joue un rôle dynamique essentiel dans de nombreux processus cellulaires tels que la transcription, la réplication, la recombinaison et la réparation de l'ADN. D'autre part, en servant de support à l'information épigénétique, la chromatine est un outil indispensable pour la modulation de l'expression des gènes.

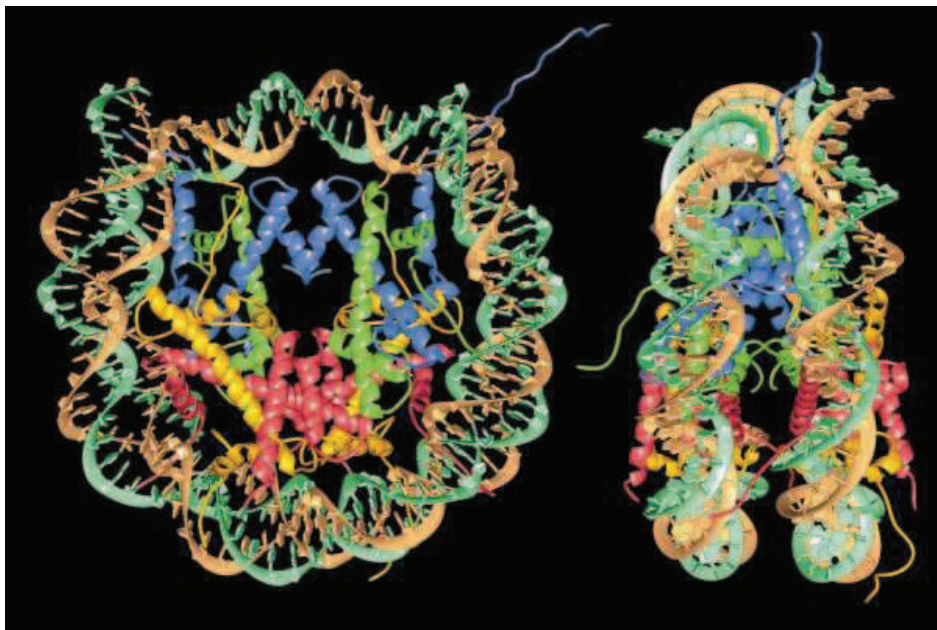
#### 1.1. Composition de la chromatine

La chromatine est un assemblage complexe d'ADN et de protéines. On distingue les protéines histones, qui participent à l'organisation structurale fondamentale de la chromatine,

et les protéines non-histones, qui comprennent une grande variété de protéines (Polycomb, MeCP2, HP1 $\alpha$ ,...) également connues sous le nom de protéines architecturales de la chromatine (CAPs, "Chromatin Architectural Proteins") et qui influencent la conformation de la chromatine (Luger and Hansen, 2005). Les histones sont de petites protéines qui comprennent cinq types: H1, H2A, H2B, H3 et H4. Très riches en acides aminés basiques, elles s'associent étroitement à l'ADN via des interactions électrostatiques fortes pour former les structures appelées nucléosomes.

## 1.2. Structure de la chromatine

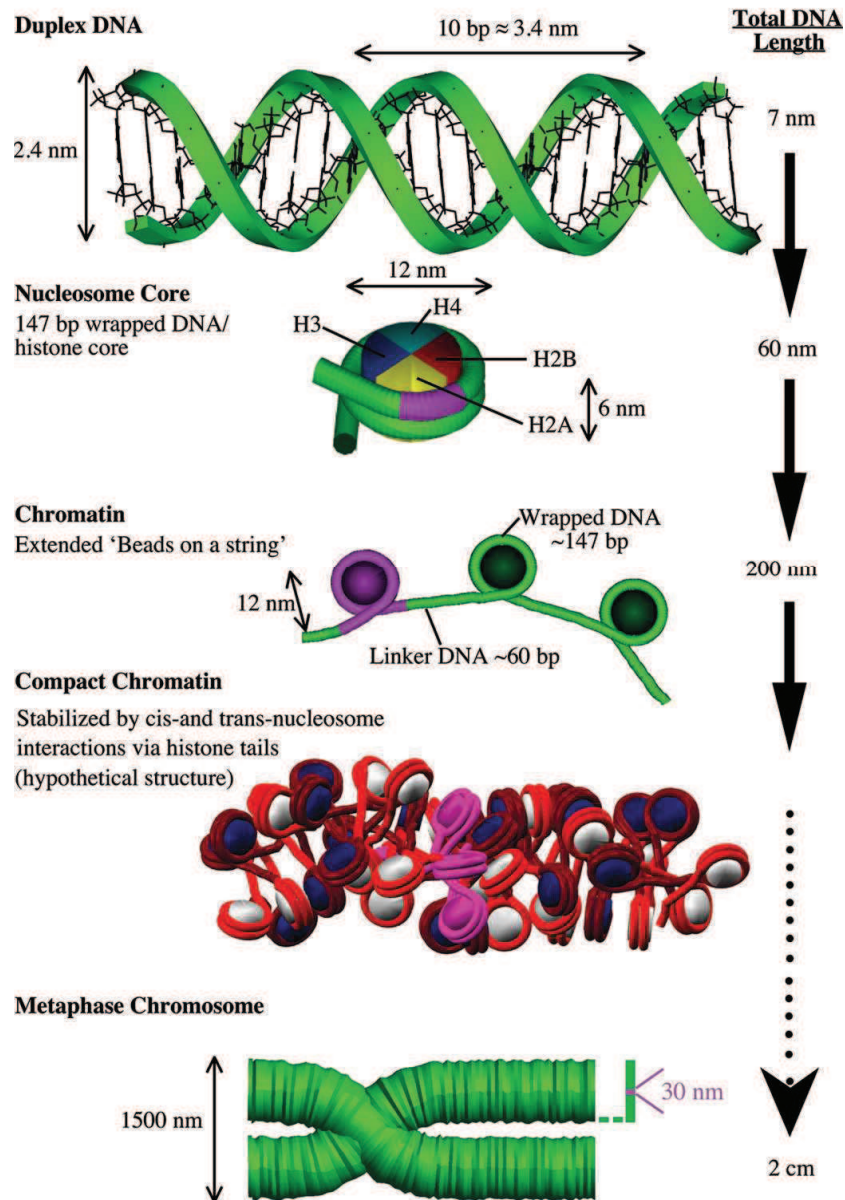
Le nucléosome, unité fondamentale de la chromatine, comprend 146 paires de bases d'ADN s'enroulant de 1,7 tours autour d'un octamère de protéines histones composé de deux copies de chacune des histones H2A, H2B, H3 et H4 (Figure 1) (Luger et al., 1997). Ces histones sont caractérisées par un domaine C-terminal globulaire, le domaine *histone-fold*, formé de trois hélices  $\alpha$  permettant la dimérisation des histones et servant de base à la formation du cœur du nucléosome. L'extrémité N-terminale des histones émerge de la surface du nucléosome et est la cible privilégiée de nombreuses modifications post-traductionnelles.



*Figure 1* : Structure cristallographique d'un nucléosome (Luger et al., 1997). La double hélice d'ADN est enroulée autour d'un octamère d'histone (bleue : H3, vert : H4, jaune : H2A, rouge : H2B). La queue N-terminale des histones émerge de la surface du nucléosome.



## DNA in the Cell

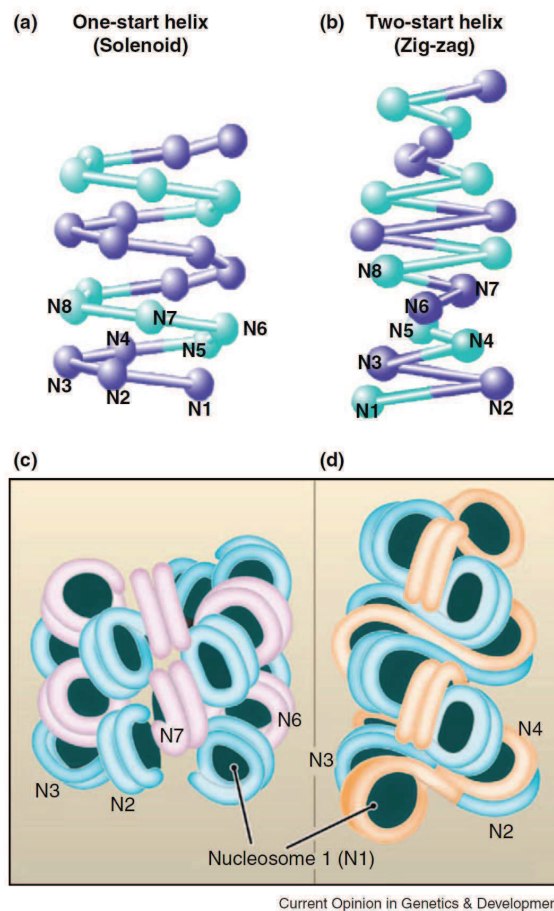


*Figure 2 : Les différents niveaux d'organisation de l'ADN dans la cellule (Schlick et al., 2012). Dans les cellules eucaryotes, l'ADN est enroulé autour d'un cœur de protéines histones pour former la fibre de chromatine. Cette fibre est présentée ici sous sa forme dépliée et sous une forme hypothétique compactée de 30 nm selon le modèle hétéromorphique en zigzag proposé par l'équipe de Schlick (Grigoryev et al., 2009). Le dernier niveau de condensation de la chromatine est atteint lorsque la fibre de chromatine se replie pour former le chromosome métaphasique. Dans chaque schéma, l'unité de l'organisation hiérarchique précédente est colorée en rose. L'échelle de longueur à droite indique le niveau de compaction impliqué.*

Au sein du noyau, le réseau de nucléosomes est ordonné en une fibre structurée en "collier de perles" de 12 nm de diamètre représentant le premier degré d'organisation de la



chromatine (Figure 2). C'est la liaison à l'ADN internucléosomique de l'histone H1, appelée histone de liaison, qui va permettre l'organisation du réseau de nucléosomes en une fibre de chromatine plus condensée de 30 nm de diamètre qui est considérée comme le second niveau d'organisation de l'ADN. Les particules résultantes, composées d'environ 167 paires de bases, de l'octamère d'histone, et d'une molécule H1 sont appelées chromatosomes (Happel and Doenecke, 2009). Le dernier niveau de condensation de la chromatine est atteint lorsque la fibre de chromatine se replie pour former le chromosome métaphasique.



*Figure 3 : Modèles d'organisation de la fibre chromatinienne de 30 nm (Li and Reinberg, 2011; Tremethick, 2007). Représentations schématiques d'une organisation en solénoïde (a et c) ou en zigzag (b et d).*

Plusieurs études ont permis de proposer deux modèles d'organisation de la fibre chromatinienne de 30 nm: le modèle en solénoïde et le modèle en zigzag (Figure 3). Grâce à une étude de microscopie électronique, l'équipe de Rhodes (Robinson et al., 2006) a permis de reconstituer une structure en solénoïde de la fibre chromatinienne dans laquelle les

nucléosomes consécutifs sont l'un à côté de l'autre, formant une hélice autour d'une cavité centrale (Figure 3b et d). Dans le modèle en zigzag proposé par l'équipe de Richmond (Dorigo et al., 2004) et conforté par l'élucidation de la structure cristallographique d'un tetranucléosome (Schalch et al., 2005), un nucléosome  $i$  est plus proche dans l'espace d'un nucléosome  $i \pm 2$ , l'ADN internucléosomique zigzagant entre les deux rangées de nucléosomes (Figure 3a et c). Plus récemment, une étude a démontré une hétérogénéité interne de la fibre compacte de 30 nm. Des ions divalents semblent promouvoir l'inflexion de l'ADN internucléosomique, provoquant la formation d'une fibre d'architecture hétéromorphe dans laquelle l'organisation en zigzag est majoritaire mais entrecoupée d'interactions de type solénoïde (Grigoryev et al., 2009).

Il semblerait donc que la chromatine puisse adopter plusieurs types de conformations, son architecture étant sensible à un grand nombre de facteurs internes et externes tels que la longueur de l'ADN internucléosomique, les variants d'histones, les modifications post-traductionnelles des histones, les conditions ioniques ou encore la liaison de protéines architecturales de la chromatine (CAPs) (Li and Reinberg, 2011; Woodcock and Ghosh, 2010). L'apparente complexité de la fibre de 30 nm reflète probablement les états de conformation variés qui sont impliqués dans la transition entre l'hétérochromatine inactive et l'euchromatine active en termes de transcription. Ainsi, la chromatine diffère dans ses propriétés structurales et dynamiques afin d'assurer ses différentes fonctions cellulaires.

L'organisation de la chromatine en une fibre de 30 nm permet de compacter l'ADN environ 50 fois. Des niveaux d'organisation plus élevés sont adoptés pour atteindre le niveau de condensation le plus élevé au sein du chromosome métaphasique. Jusqu'à présent, les niveaux de compaction au delà de la fibre de 30 nm sont très peu caractérisés. Le degré de compaction de l'ADN module son accessibilité et en conséquence l'expression des gènes qu'il porte.

### 1.3. Euchromatine et hétérochromatine

Les termes d'hétérochromatine et d'euchromatine désignent des états de compaction et le potentiel de transcription de régions de la chromatine (Woodcock and Ghosh, 2010). Ainsi, l'hétérochromatine et l'euchromatine correspondent respectivement aux états condensés et

décondensés de la chromatine, la transcription étant restreinte en grande partie à l'euchromatine. L'hétérochromatine peut être subdivisée en hétérochromatine constitutive ou hétérochromatine facultative. L'hétérochromatine constitutive est toujours compacte, comprend peu de gènes et est formée principalement de séquences répétées. De la même manière que l'hétérochromatine constitutive, l'hétérochromatine facultative est transcriptionnellement inactive mais conserve cependant le potentiel de se convertir en euchromatine pour permettre la transcription dans certains contextes: temporel (stade de développement, stade spécifique du cycle cellulaire), spatial (modification de la localisation nucléaire grâce à des facteurs ou des signaux exogènes) ou parental/héréditaire (expression monoallélique de gènes soumis à l'empreinte génomique parentale) (Trojer and Reinberg, 2007). Pendant l'embryogenèse par exemple, la quantité d'hétérochromatine facultative augmente lorsque les gènes superflus sont progressivement réprimés, jusqu'à maturation des cellules qui expriment alors uniquement les gènes tissus-spécifiques (Woodcock and Ghosh, 2010). L'inverse est observé lorsque des cellules différenciées sont reprogrammées pour se reconvertir en cellules souches. Ces événements sont typiquement accompagnés de profonds changements épigénétiques (Woodcock and Ghosh, 2010).

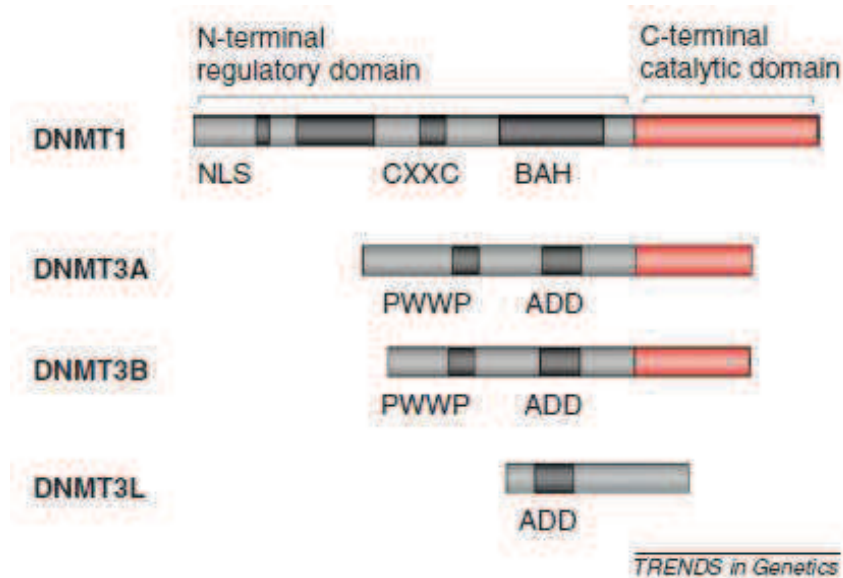
## 2. La méthylation de l'ADN

La méthylation de l'ADN étant impliquée dans la régulation des gènes, la formation et la maintenance de la chromatine, le développement, et d'autres processus fondamentaux, elle constitue un élément clé de la régulation épigénétique de l'expression des gènes.

### 2.1. Méthylation par les méthyltransférases de l'ADN

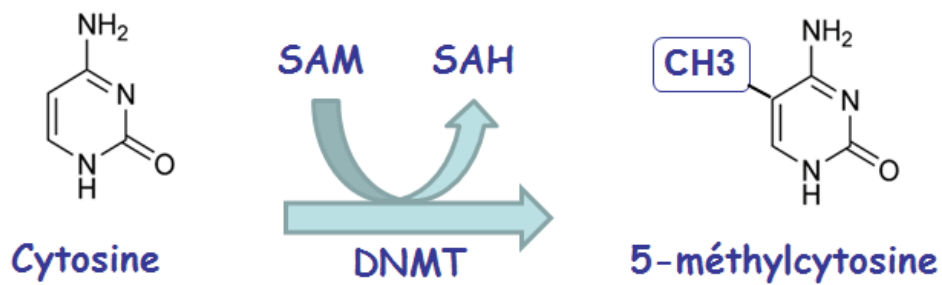
La méthylation de l'ADN est une modification post-réplivative qui, chez les vertébrés, se produit exclusivement sur le carbone 5 du cycle pyrimidine de résidus cytosines. L'addition covalente d'un groupement méthyl sur une cytosine est catalysée par les méthyltransférases de l'ADN (Dnmts) (Rottach et al., 2009). Fondé sur des critères d'homologie de séquences, cette famille était à l'origine composée de cinq membres: Dnmt1, Dnmt3a, Dnmt3b, Dnmt3L et Dnmt2 (Figure 4) (Fernandez et al., 2012b). Cependant, il fut démontré plus tard que la

Dnmt2, présentant une activité méthyltransférase de l'ADN très faible, était préférentiellement impliquée dans la méthylation de l'ARNt (Goll et al., 2006; Hermann et al., 2004).

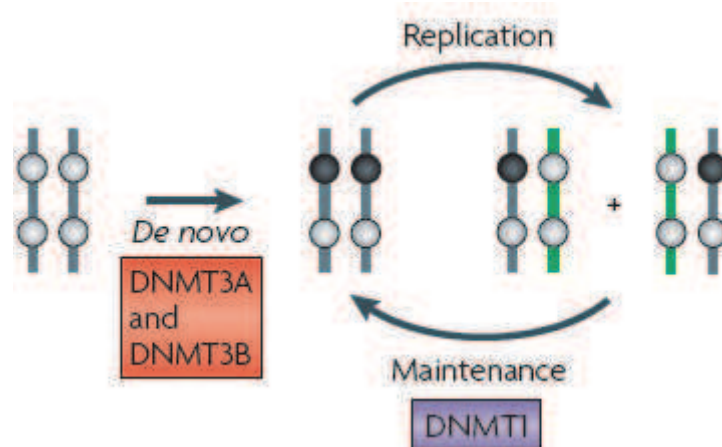


*Figure 4* : Représentation schématique des protéines de mammifères *Dnmt1*, *Dnmt3a*, *Dnmt3b*, et *Dnmt3L* (Fernandez et al., 2012b). ADD: domaine d'interaction aux histones; BAH: "Bromo-adjacent homology domain"; CXXC: domaine contenant un motif de liaison du  $Zn^{2+}$  riche en cystéines; NLS: signal de localisation nucléaire ("nuclear localization signal"); PWWP: domaine contenant un motif proline-tryptophane-tryptophane-proline hautement conservé.

Les Dnmts de mammifères possèdent deux domaines principaux : une région N-terminale de taille variable qui possède des fonctions de régulation, et une région C-terminale catalytique (Figure 4) (Jurkowska et al., 2011). La partie N-terminale guide la localisation nucléaire des enzymes et permet des interactions avec d'autres protéines, l'ADN et la chromatine. Le domaine catalytique C-terminal est hautement conservé et également présent au sein des (cytosine-5)-méthyltransférases de l'ADN des procaryotes (Cheng, 1995; Goll and Bestor, 2005). Ainsi, toutes ces enzymes adoptent le même mécanisme catalytique: après reconnaissance du substrat, la cytosine ciblée est basculée hors de la double hélice d'ADN, l'enzyme formant un complexe covalent avec le carbone 6 de la cytosine. Le transfert du groupement méthyle du cofacteur S-adénosylméthionine (AdoMet, SAM) sur la position C5 activée permet la formation d'une 5-méthylcytosine et la libération de la S-adénosylhomocystéine (AdoHcy, SAH) (Figure 5) (Rottach et al., 2009; Stresemann and Lyko, 2008).



*Figure 5 : Méthylation de l'ADN sur le carbone 5 d'un résidu cytosine. Les méthyltransférases de l'ADN catalysent le transfert d'un groupement méthyle du cofacteur S-adenosylméthionine (SAM) sur le carbone 5 d'un résidu cytosine. Il en résulte la formation de la 5-méthylcytosine et la libération de la S-adenosylhomocystéine (SAH).*



*Figure 6 : Modèle actuellement proposé pour l'établissement et la maintenance des profils de méthylation de l'ADN (Jones and Liang, 2009). Les profils de méthylation de l'ADN sont initialement établis dans les cellules germinales et embryonnaires par les de novo méthyltransférases de l'ADN, Dnmt3a et Dnmt3b. Par la suite, les profils de méthylation (les cercles noirs représentent les cytosines méthylées et les cercles gris les cytosines non-méthylées) sont transmis après la réplication de l'ADN principalement par la Dnmt1 qui se lie préférentiellement aux sites héli-méthylés générés par la synthèse de l'ADN (le brin "fille" est présenté en vert). L'enzyme copie donc un patron de méthylation présent sur le brin d'ADN "parental".*

Les mécanismes par lesquels les profils de méthylation sont perpétués de cellule en cellule au sein de l'organisme ont été suggérés pour la première fois par Riggs et Holliday en 1975 (Holliday and Pugh, 1975; Riggs, 1975). Ces auteurs ont proposé l'existence de profils de méthylation tissu-spécifiques pouvant moduler l'expression de gènes en influençant la

liaison de facteurs de transcription et d'autres protéines à l'ADN (Jones and Liang, 2009). Dans le modèle proposé, les profils de méthylation de l'ADN sont initialement établis dans les cellules germinales et embryonnaires par les *de novo* Dnmts (Dnmt3a et Dnmt3b). Ce patron de méthylation est perpétué de cellule en cellule durant le reste de la vie de l'animal, avec néanmoins quelques variations tissus-spécifiques. Après chaque phase de réplication, l'ADN hémi-méthylé nouvellement formé possède le brin parental portant les marques de méthylation et le brin nouvellement synthétisé dépourvu de méthylations. Pour assurer la pérennité des profils de méthylation, la Dnmt de maintenance (Dnmt1), spécifique de l'ADN hémi-méthylé, va méthyler le brin "fille" en copiant le patron de méthylation présent sur le brin "parental" (Figure 6) (Jurkowska et al., 2011).

Actuellement, il est donc bien établi que la Dnmt1 est une méthyltransférase de maintenance jouant un rôle majeur dans le maintien des profils de méthylation au cours des divisions cellulaires (Jones and Liang, 2009; Jurkowska et al., 2011; Rottach et al., 2009). Contrairement aux autres Dnmts, cette enzyme possède une préférence marquée pour l'ADN hémi-méthylé par rapport à l'ADN non-méthylé (Fatemi et al., 2001; Goyal et al., 2006; Hermann et al., 2004). De plus, la Dnmt1 est une protéine ubiquitaire, de loin la plus abondante des Dnmts dans les cellules somatiques proliférantes. Ainsi, elle est principalement transcrite lors de la phase S du cycle cellulaire (Leonhardt et al., 1992; Robertson et al., 2000b), lorsque la méthylation des sites hémi-méthylés nouvellement synthétisés est le plus nécessaire. Son interaction avec PCNA ("proliferating cell nuclear antigen"), la pince de réplication de l'ADN, lui permet de se localiser au niveau de la fourche de réplication en augmentant son affinité pour l'ADN (Chuang et al., 1997; Easwaran et al., 2004).

La délétion du gène de la Dnmt1 a montré que cette enzyme était responsable de l'essentiel de la méthylation, et qu'elle était nécessaire au développement embryonnaire (Li et al., 1992). Ainsi, en plus de son rôle primordial dans la maintenance des profils de méthylation, l'activité de la Dnmt1 semble également requise pour la méthylation *de novo* (Arand et al., 2012; Egger et al., 2006; Feltus et al., 2003; Jair et al., 2006). Dans cette fonction, la Dnmt1 assisterait la Dnmt3a et la Dnmt3b en se servant des sites hémi-méthylés générés par les enzymes Dnmt3 comme substrat (Fatemi et al., 2002) (Jurkowska et al., 2011).



La famille Dnmt3 des Dnmts comprend trois membres: Dnmt3a, Dnmt3b et Dnmt3L. Par opposition à la Dnmt1, la Dnmt3a et la Dnmt3b sont considérées comme les *de novo* méthyltransférases car elles sont en grande partie responsables de l'établissement des profils de méthylation lors du développement embryonnaire (Okano et al., 1999). En effet, alors que les *de novo* Dnmts sont largement exprimées dans les cellules embryonnaires, leur niveau d'expression est significativement diminué dans la plupart des tissus adultes lorsque les cellules sont différenciées (Bestor, 2000). De plus, de même que les Dnmts procaryotes, les *de novo* Dnmts présentent des affinités équivalentes pour l'ADN hémi-méthylé ou non-méthylé (Gowher and Jeltsch, 2001; Okano et al., 1998).

Malgré son homologie de séquence avec les enzymes Dnmt3a et Dnmt3b, le troisième membre de la famille Dnmt3, Dnmt3L, est enzymatiquement inactif car dépourvu de motif catalytique (Figure 4) (Bourc'his et al., 2001; Fernandez et al., 2012b; Jurkowska et al., 2011; Rottach et al., 2009). Ainsi, il est incapable de lier le cofacteur SAM et présente une affinité faible pour l'ADN (Gowher et al., 2005; Jurkowska et al., 2011; Kareta et al., 2006). Cependant, Dnmt3L interagit avec Dnmt3a et Dnmt3b en stimulant leur activité enzymatique (Chedin et al., 2002; Chen et al., 2005; Gowher et al., 2005; Hata et al., 2002; Kareta et al., 2006; Suetake et al., 2004) vraisemblablement par induction d'un changement de conformation qui faciliterait la liaison de l'ADN et du cofacteur ainsi que la catalyse (Gowher et al., 2005; Jia et al., 2007; Jurkowska et al., 2011). Ainsi, malgré son absence d'activité enzymatique, la Dnmt3L joue le rôle de régulateur positif des *de novo* Dnmts et est essentielle pour l'établissement de l'empreinte génomique parentale et la méthylation d'éléments rétrotransposables dans les cellules mâles germinatives (Bourc'his and Bestor, 2004; Bourc'his et al., 2001; Hata et al., 2002; Jurkowska et al., 2011; Margot et al., 2003; Rottach et al., 2009).

La Dnmt3a, associée à la Dnmt3L, est nécessaire à l'établissement des profils de méthylation lors de la gamétogenèse, contrairement à la Dnmt3b qui est accessoire lors de ce processus (Jurkowska et al., 2011). Cependant, la Dnmt3a et la Dnmt3b sont toutes deux impliquées dans la méthylation des divers types de séquences répétées, la Dnmt3b étant responsable de la méthylation des séquences satellites des centromères (Jurkowska et al., 2011). Bien que la famille des Dnmt3s joue un rôle essentiellement dans la méthylation *de novo*, ces enzymes semblent également requises pour la maintenance fidèle des profils de méthylation au niveau de l'hétérochromatine dans les cellules somatiques et les cellules souches embryonnaires (Arand et al., 2012; Chen et al., 2003; Dodge et al., 2005; Fernandez

et al., 2012b; Jeong et al., 2009; Jurkowska et al., 2011 ; Kim et al., 2002; Liang et al., 2002; Rottach et al., 2009).

## 2.2. Les sites de méthylation de l'ADN

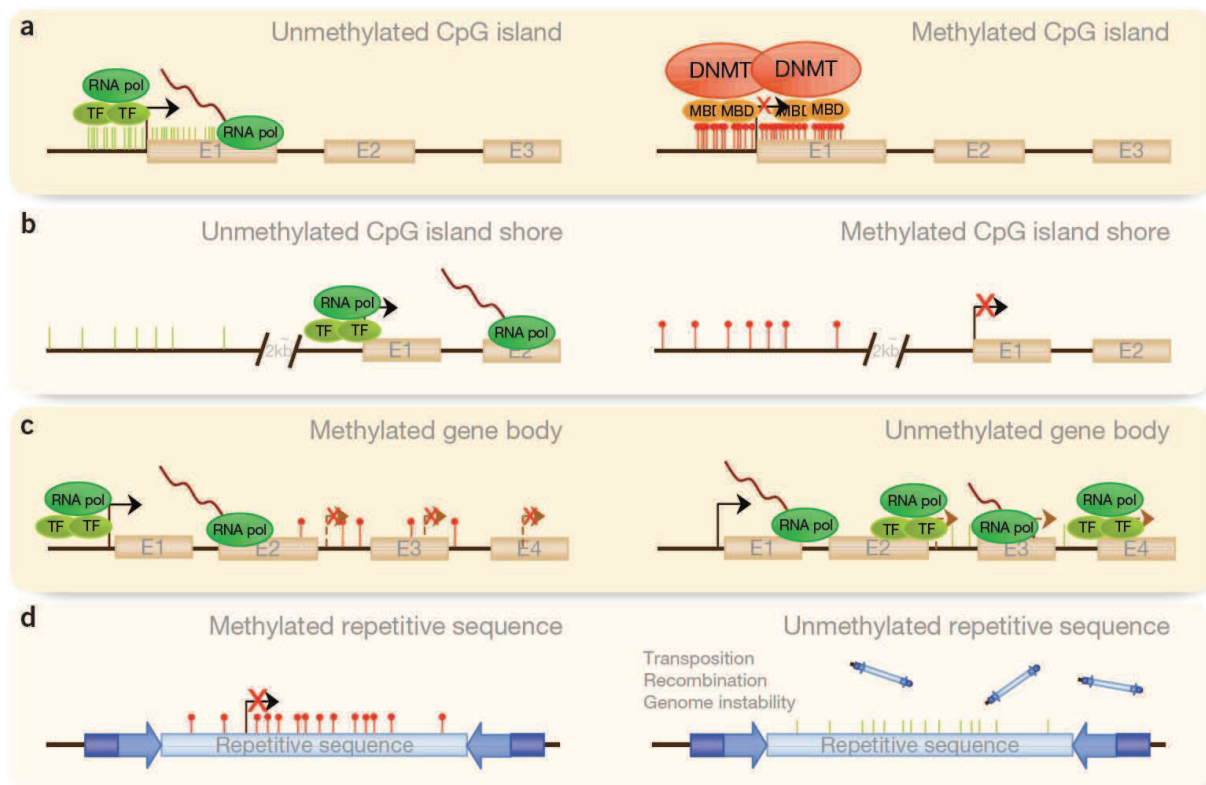
Chez les mammifères, la méthylation de l'ADN se produit principalement sur les résidus cytosines de dinucléotides palindromiques 5'-CG-3', appelés dinucléotides CpG, et occasionnellement dans un contexte non-CpG au sein de nucléotides 5'-CHG-3' et 5'-CHH-3' où H = A, C ou T (Lister et al., 2009; Ramsahoye et al., 2000). De façon générale, les dinucléotides CpG ne sont pas très abondants dans le génome des mammifères du fait de la désamination spontanée de la 5-méthylcytosine en thymidine (Illingworth and Bird, 2009; Jones, 2012; Jurkowska et al., 2011; Lander et al., 2001; Pfeifer et al., 2000). Les sites CpG sont répartis de façon hétérogène dans le génome avec une sous-représentation générale mais un enrichissement dans les séquences répétées et des séquences courtes d'environ 1 kb, appelées îlots CpG (Antequera, 2003), qui recouvrent partiellement les régions promotrices et/ou le premier exon de 60 % de l'ensemble des gènes constitutifs humains (Fernandez et al., 2012b; Illingworth and Bird, 2009; Saxonov et al., 2006; Weber et al., 2007). La localisation des îlots CpG au niveau des promoteurs et à proximité du site d'initiation de la transcription, révèle l'importance de la méthylation de l'ADN dans la régulation de l'expression des gènes. En effet, l'hyperméthylation des îlots CpG au niveau des régions promotrices des gènes est corrélée avec l'inhibition de l'expression des gènes (Figure 7) (Fernandez et al., 2012b).

Dans le génome humain, approximativement 70 % de l'ensemble des dinucléotides CpG sont méthylés (Bird, 2002; Ehrlich et al., 1982), les îlots CpG échappant en général à la méthylation (Figure 7) (Sasai and Defossez, 2009; Wong et al., 2007). En effet, dans les cellules somatiques, la plupart des îlots CpG restent non-méthylés indépendamment de l'état d'expression des gènes avec comme exception les gènes impliqués dans la répression à long terme ("long-term silencing") tels que les gènes sujets à l'empreinte génomique parentale, les gènes du chromosome X inactivé, les gènes spécifiques des cellules germinales, ou les gènes impliqués dans la répression tissu-spécifique (Fernandez et al., 2012b; Illingworth and Bird, 2009; Jones, 2012; Jones and Liang, 2009). Inversement, les sites CpG localisés dans les régions des séquences répétées ou les transposons, les régions intergéniques, et les "gene bodies", qui sont des régions généralement moins denses en sites CpG, sont habituellement



hyperméthylées (Figure 7) (Fernandez et al., 2012b; Hodges et al., 2009; Jones, 2012; Laurent et al., 2010; Lister et al., 2009). La méthylation au niveau de ces sites de régulation influence la liaison et les fonctions de protéines de régulation. Contrairement aux promoteurs, la méthylation des "gene bodies" favorise la transcription en évitant l'initiation de transcriptions aberrantes. Dans le cas des séquences répétées, leur méthylation préviendrait l'instabilité chromosomique, les translocations, et la disruption génique par réactivation de séquences endoparasites (Portela and Esteller, 2010). Bien que ces mécanismes ne soient pas encore bien compris, on suppose qu'ils jouent un rôle essentiel dans le développement, la différenciation et donc la viabilité cellulaire (Jones, 2012). Ainsi, par comparaison avec les régions promotrices des gènes contenant des îlots CpG, qui sont généralement non-méthylées et qui semblent assurer la répression des gènes à long-terme, les profils de méthylation des régions du génome possédant une faible densité de sites CpG semblent plus intéressants (Jones, 2012). En effet, c'est le profil de méthylation de ces régions, et plus particulièrement des régions promotrices des gènes possédant une faible densité de sites CpG, qui est généralement cellule et/ou tissu-dépendant (Calvanese et al., 2012; Fernandez et al., 2012a; Fernandez et al., 2012b). De plus, il a récemment été démontré que la plupart des méthylations tissus-spécifiques de l'ADN n'a pas lieu au niveau des îlots CpG mais au niveau de région appelées "CpG island shores". Ces régions régulatrices contiennent une faible densité de sites CpG et se situent en amont de plus de 2 kb des sites d'initiation de la transcription (Figure 7). Ces plages sont enrichies en régions de méthylation différentielle appelées "T-DMRs" ("tissue specific differentially methylated region") et semblent jouer un rôle important dans la différenciation cellulaire (Irizarry et al., 2009).

Ainsi, la méthylation des sites CpG ne s'effectue que sur certains sites ciblés et induit la formation de profils de méthylation tissus- et cellules-spécifiques. Leur caractérisation permet d'établir une carte des cytosines méthylées, ou méthylome, qui constitue l'empreinte spécifique d'une cellule, d'un tissu ou d'un organisme.



*Figure 7 : Profils de méthylation de l'ADN (Portela and Esteller, 2010). La méthylation de l'ADN peut se produire au niveau de différentes régions du génome. L'altération de ces profils de méthylation est liée à l'apparition de diverses maladies. Le cas de figure normal est illustré dans la colonne de gauche et les altérations de ces profils sont présentées à droite. (a) Les îlots CpG au niveau des promoteurs de gènes sont habituellement non-méthylés, autorisant la transcription. Une hyperméthylation aberrante entraîne l'inactivation de la transcription. (b) Le même profil est observé lorsque l'on étudie les "CpG island shores" qui sont localisés en amont de 2 kb des sites d'initiation de la transcription. (c) Cependant, lorsque la méthylation se produit au niveau des "gene bodies", la transcription est favorisée en évitant l'initiation de transcriptions aberrantes. Lors d'altérations, les "gene bodies" ont tendance à être déméthylés, autorisant l'initiation de la transcription au niveau de plusieurs sites incorrects. (d) Enfin, les séquences répétées sont hyperméthylées, prévenant l'instabilité chromosomique, les translocations et la disruption génique par réactivation des séquences endoparasites.*

La méthylation de résidus cytosines est occasionnellement observée dans un contexte non-CpG au sein de nucléotides 5'-CHG-3' et 5'-CHH-3' où H = A, C ou T (Lister et al., 2009; Ramsahoye et al., 2000). La présence de ces méthylation non-CG est caractéristique des cellules souches embryonnaires (Laurent et al., 2010; Lister et al., 2009) et disparaît dès la différenciation cellulaire. Ces sites de méthylation semblent être la cible préférentielle des Dnmt3s (Aoki et al., 2001; Arand et al., 2012; Gowher and Jeltsch, 2001; Lister et al., 2009; Ramsahoye et al., 2000). Ainsi, l'absence des sites de méthylation 5'-CHG-3' et 5'-CHH-3' dans les cellules différenciées coïncide avec la diminution significative des Dnmts *de novo*.

De plus, une étude structurale de la méthyltransférase Dnmt3a et de sa protéine partenaire Dnmt3L a démontré que deux copies de chaque Dnmt peuvent former un hétérotétramère contenant deux sites actifs séparés d'une longueur équivalente à 8-10 nucléotides dans une hélice d'ADN (Ferguson-Smith and Greally, 2007; Jia et al., 2007). Dans le génome humain, la distribution régulière des sites non-CG à 8-10 paires de bases de distance, suggère que la Dnmt3a pourrait être responsable de la méthylation de ces sites (Lister et al., 2009).

### 2.3. Les modules de reconnaissance des CpG méthylés

De manière générale, lorsqu'elle a lieu au niveau des promoteurs, la méthylation de l'ADN est associée à une inhibition de la transcription. Deux mécanismes complémentaires semblent être impliqués dans ce phénomène (Parry and Clarke, 2011; Rottach et al., 2009; Sasai and Defossez, 2009). D'une part, lorsqu'elles sont présentes au niveau de leur site de reconnaissance, les marques de méthylation peuvent directement prévenir la liaison de facteurs de transcription (Bell and Felsenfeld, 2000; Bird, 2002; Hark et al., 2000). D'autre part, la méthylation de l'ADN crée des sites de liaison spécifiquement reconnus par les protéines se liant à l'ADN méthylé, les MBPs ("methyl-binding protein"). Ces protéines recrutent des complexes de remodelage de la chromatine qui contiennent des enzymes de modifications post-traductionnelles des histones, des enzymes responsables de la méthylation de l'ADN et des corépresseurs transcriptionnels, qui concourent à l'établissement de l'hétérochromatine. Cette zone devient alors inaccessible aux complexes de transcription, empêchant ainsi l'expression des gènes (Fuks, 2005).

Chez les mammifères, les sites CpG méthylés sont reconnus par au moins trois familles de protéines MBPs (Figure 8): la famille de protéines contenant un domaine de reconnaissance des CpG méthylé MBD ("methyl-CpG binding domain"), la famille de protéines contenant des doigts de zinc, et la famille de protéines contenant le domaine SRA ("SET and RING Associated domain") (Parry and Clarke, 2011; Rottach et al., 2009; Sasai and Defossez, 2009). Actuellement, grâce à des homologies de séquences, 16 MBPs ont été identifiées et réparties dans ces trois familles (Parry and Clarke, 2011). Chaque famille emprunte un mécanisme différent pour interagir avec l'ADN méthylé ou ses composants afin de réguler l'expression des gènes et de maintenir ou altérer l'architecture de l'ADN.

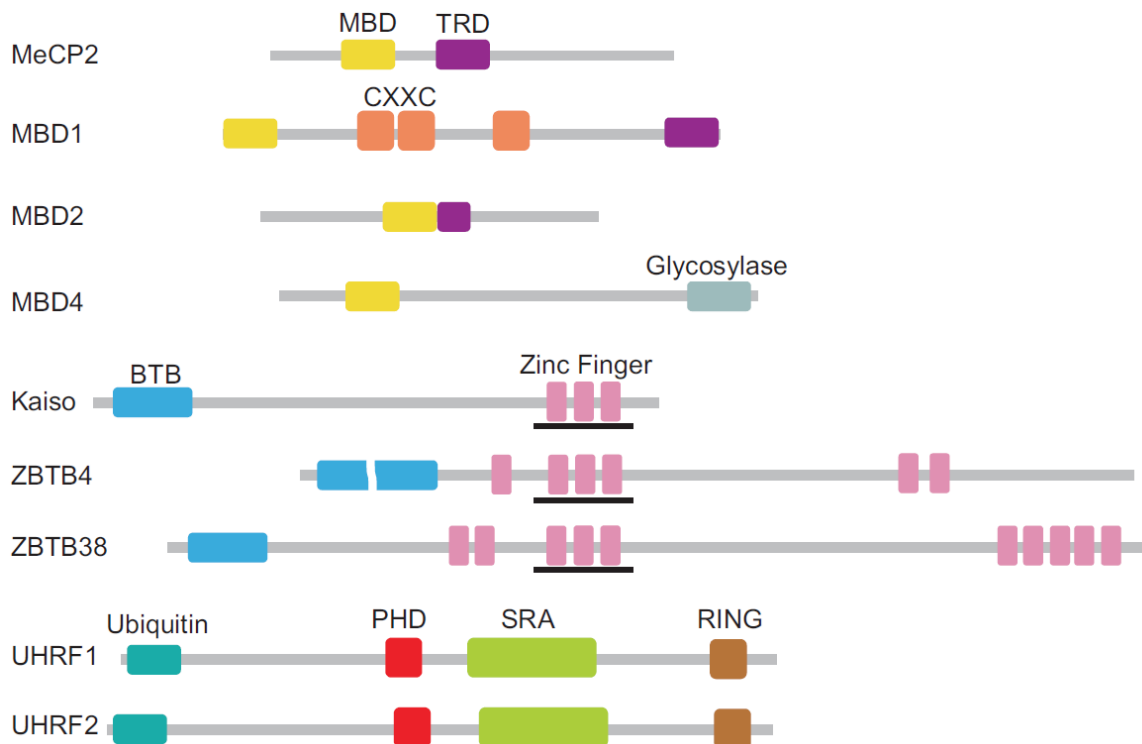
Basé sur la présence d'autres domaines, la superfamille des protéines MBDs est subdivisée en trois sous-groupes: les méthyltransférases des histones (HMT-MBD, "histone methyltransferases methyl-binding domain"), les protéines MeCP2, et les acétyltransférases des histones (HAT-MBD, "histone acetyltransferases methyl-binding domain") (Parry and Clarke, 2011). Bien que tous les membres de cette superfamille contiennent le domaine MBD, elles n'interagissent pas forcément de manière directe avec les sites CpG méthylés.

Les HMT-MBD comprennent deux membres: SETDB1 (aussi connu sous le nom de ESET) (Schultz et al., 2002) et SETDB2 (aussi connu sous le nom de CLLD8) (Falandry et al., 2010). Ces deux protéines contiennent le domaine SET responsable de l'activité méthyltransférase qui permet la méthylation spécifique de l'histone H3 au niveau de la lysine 9 (H3K9) et conduit à la répression de la transcription grâce à la formation de l'hétérochromatine. Bien que ces deux protéines contiennent le domaine MBD, leur liaison à l'ADN méthylé reste à confirmer (Hashimoto et al., 2010).

La famille MeCP2 représente le groupe le plus important des protéines MBDs. Il comprend la protéine MeCP2, membre fondateur de la famille, ainsi que six autres membres, MBD1 à MBD6 (Parry and Clarke, 2011). Mis à part les protéines MBD3, MBD5 et MBD6, toutes les protéines de la famille MeCP2 se lient spécifiquement aux CpG méthylées (Hendrich and Tweedie, 2003). De plus, la plupart de ces protéines (excepté MBD4, MBD5 et MBD6) forment des complexes avec des déacétylases des histones (HDACs) et des complexes de remodelage de la chromatine tels que MeCP1 et NuRD, associés à la répression transcriptionnelle (Hashimoto et al., 2010; Rottach et al., 2009).

Les HAT-MBD comprennent deux membres: BAZ2A (aussi connu sous le nom de TIP5) et BAZ2B. Ces deux protéines contiennent les domaines MBD, DDT ("DNA binding homeobox and Different Transcription factors"), PHD et un bromodomaine, consécutivement arrangés dans le même ordre (Hung and Shen, 2003). Le domaine MBD de ces protéines diffère de quelques résidus par rapport aux domaines des autres membres de la superfamille de protéines MBDs et ces deux protéines ne semblent pas avoir d'affinité préférentielle pour l'ADN méthylé. BAZ2A est un composé majeur du complexe de remodelage nucléolaire (NoRC, "nucleolar remodeling complex") qui induit la répression de la transcription de gènes ribosomaux via la déacétylation des histones (Zhou et al., 2002).

Trois protéines sont connues pour reconnaître les CpG méthylés via des doigts de zinc: Kaiso, ZBTB4 et ZBTB38. Ces protéines partagent le même motif composé d'un tandem de trois doigts de zinc situé dans la partie C-terminale des protéines. La protéine Kaiso est capable de reconnaître une paire de dinucléotides CpG méthylés (Prokhortchouk et al., 2001). Les deux autres protéines ZBTB4 et ZBTB38, quant à elles, ne nécessitent qu'un seul dinucléotide CpG méthylé pour interagir avec l'ADN (Filion et al., 2006). Des analyses biochimiques ont révélé l'interaction directe de Kaiso avec le complexe de répression NCoR ("nuclear receptor corepressor"), qui contient également HDAC1 et des protéines de remodelage ("remodeling activities"), reliant à nouveau la méthylation de séquences d'ADN à un état hautement condensé de la chromatine (Rottach et al., 2009).



*Figure 8 : Représentation schématique des principales protéines humaines MBPs liant l'ADN méthylé (Sasai and Defossez, 2009). Trois domaines protéiques sont identifiés comme des modules de reconnaissance de l'ADN méthylé: le domaine MBD, les doigts de zinc, et le domaine SRA. Les protéines MBD contiennent MBD ainsi que le domaine de trans-répression (TRD, "trans-repression domain"), les doigts de zinc CXXC ou le domaine glycosylase. Les doigts de zinc sont représentés par des rectangles roses, et ceux responsables de la liaison à l'ADN méthylé sont soulignés. Les protéines contenant des doigts de zinc présentent également un domaine BTB et plus de 7 doigts de zinc additionnels. Le domaine BTB de la protéine ZBTB4 est interrompu par l'insertion d'une séquence riche en sérines. Les protéines UHRF1 et UHRF2 contiennent le domaine SRA, un domaine Ubiquitin-like (UBL), un domaine PHD ("Plant Homeo Domain") et un domaine RING ("Really Interesting New Gene").*

La troisième classe de MBPs est composée des protéines UHRF1 (aussi appelée Np95, ou ICBP90) et UHRF2 (aussi appelée NIRF) (Bronner et al., 2007a). Les deux membres de cette famille sont les seules protéines de mammifère à contenir le domaine SRA ("SET and RING Associated domain") qui est responsable de leur liaison à l'ADN au niveau des sites CpG méthylés. De nombreuses observations suggèrent un rôle de UHRF1 et UHRF2 dans la connexion entre la méthylation de l'ADN et les modifications des histones. D'autre part, la protéine UHRF1, qui reconnaît préférentiellement l'ADN hémi-méthylé, jouerait un rôle important dans le maintien des profils de méthylation de l'ADN en dirigeant la Dnmt1 au niveau des sites hémi-méthylés (Arita et al., 2008; Avvakumov et al., 2008; Bostick et al., 2007; Hashimoto et al., 2008; Sharif et al., 2007). Les fonctions de cette famille de protéines, et plus particulièrement celles de la protéine UHRF1, seront détaillées plus loin (voir p. 55).

## 2.4. La méthylation *de novo*

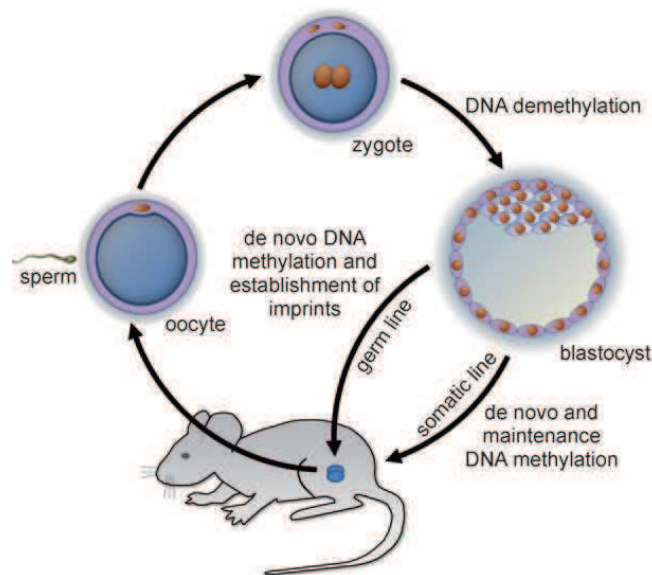
Chez les mammifères, le degré de méthylation de l'ADN est régulé de manière dynamique au cours du développement (Figure 9). Après la fécondation, le zygote subit une reprogrammation intensive pour établir la totipotence dans l'embryon (Morgan et al., 2005; Reik et al., 2001). Les empreintes épigénétiques qui sont conservées lors des divisions cellulaires sont des composants clés de l'identité cellulaire et participent au maintien des cellules à l'état différencié. A ce stade du développement, les marques épigénétiques préexistantes doivent donc être réinitialisées afin d'établir la totipotence du zygote qui donnera lieu à toutes les cellules de l'embryon. Ce processus est appelé reprogrammation épigénétique (Feng et al., 2010; Smallwood and Kelsey, 2012). Ainsi, dans le zygote, une déméthylation quasiment complète du génome est observée. Certaines régions du génome appelées "gDMRs" ("germline differentially methylated regions") échappent toutefois à cette vague de déméthylation. Ces régions de méthylation différentielle (DMR) spécifiques des cellules germinales mâles et femelles conservent leurs méthylations tout au long du développement et contrôlent l'expression monoallélique des gènes soumis à l'empreinte génomique parentale (Smallwood and Kelsey, 2012).

Pendant les premiers stades du développement embryonnaire, lors de la nidation, les profils de méthylation de l'ADN sont nouvellement établis par les enzymes Dnmt3s et sont maintenus pendant toute la durée de vie de l'organisme dans la lignée somatique. Au cours du



développement de la lignée germinale, les cellules germinales primordiales issues des cellules embryonnaires subissent une nouvelle vague de déméthylation afin de pouvoir acquérir de nouvelles marques épigénétiques spécifiques. Les marques de méthylation initialement établies dans les cellules germinales mâles sont rapidement effacées par un mécanisme de déméthylation actif impliquant probablement l'oxydation par les protéines Tet (Gu et al., 2011). Les marques de méthylation maternelles seraient perdues de manière passive par un déficit de la méthylation de maintenance lors de la réplication, induisant la perte progressive des méthylations à chaque division cellulaire (Smallwood and Kelsey, 2012). Après la détermination du sexe de l'embryon, de nouveaux profils de méthylation spécifiques de l'ovocyte ou du spermatozoïde sont établis et des séquences variées sont alors la cible de la méthylation *de novo*. Ainsi, selon le type de gamète, des îlots CpG (Smallwood et al., 2011) et des régions "gDMRs" sont spécifiquement méthylés (Ferguson-Smith, 2011; Suetake et al., 2004).

Des phases de déméthylation et de méthylation créent donc un cycle épigénétique chez les mammifères et conduisent à la génération de profils de méthylation distincts qui est une caractéristique majeure du génome de mammifère (Jurkowska et al., 2011; Smallwood and Kelsey, 2012).



*Figure 9 : Dynamique de la méthylation de l'ADN pendant le développement des mammifères (Jurkowska et al., 2011). Après la fécondation, le zygote subit une déméthylation globale de l'ADN afin d'établir la totipotence dans l'embryon. Les profils de méthylation de l'ADN sont établis de novo dans la lignée somatique lors de la nidation et sont maintenus pendant toute la durée de vie de l'organisme. Lors du développement des cellules germinales, une nouvelle vague de méthylation de novo se produit et les empreintes de méthylation sont établies de manière spécifique selon le sexe.*

Chez les mammifères, il est donc généralement admis que le profil de méthylation initial est généré pendant les premiers stades du développement embryonnaire par les *de novo* méthyltransférases Dnmt3a, Dnmt3b et Dnmt3L. Ces trois enzymes, exprimées à la fois dans les cellules germinales mâles et femelles (Lucifero et al., 2007; Lucifero et al., 2004), jouent également un rôle central dans l'établissement des profils de méthylation lors de la gamétogenèse (Bourc'his et al., 2001; Kaneda et al., 2010; Kaneda et al., 2004), la Dnmt3b n'étant cependant nécessaire que dans les cellules germinales mâles (Kato et al., 2007; Smallwood and Kelsey, 2012).

Au cours du développement, la méthylation ciblée de certains sites CpG est impliquée dans la différenciation cellulaire et l'établissement de l'empreinte génomique parentale. Les mécanismes par lesquels les Dnmts choisissent leur cibles et induisent la méthylation de seulement certains sites CpG sont encore mal compris.

Dans les cellules germinales, des séquences d'ADN caractéristiques pourraient permettre le recrutement des Dnmts au niveau des zones à méthyler. En effet, les sites non-CpG trouvés dans les cellules souches embryonnaires, principalement dans un contexte CA (Laurent et al., 2010; Lister et al., 2009), représentent les cibles privilégiées des Dnmts *de novo*, Dnmt3a et Dnmt3b (Aoki et al., 2001; Arand et al., 2012; Gowher and Jeltsch, 2001; Ramsahoye et al., 2000), suggérant que la distribution des méthylations pourrait refléter la préférence des enzymes Dnmt3s pour ces séquences spécifiques. D'autre part, il a aussi été proposé que l'espace entre les sites CpG au sein des îlots CpG puisse déterminer l'orientation des Dnmt3s au niveau de certaines séquences. En effet, le complexe formé par Dnmt3a et Dnmt3L favorise la méthylation des sites qui sont régulièrement distribués à 8-10 nucléotides de distance au sein de l'hélice d'ADN (Ferguson-Smith and Greally, 2007; Jia et al., 2007; Lister et al., 2009). Supportant cette idée, la périodicité de 8-10 paires de bases entre les sites méthylés a été observée dans divers contextes, notamment dans les régions "gDMR" et les méthylations CpG et non-CpG (Jia et al., 2007; Lister et al., 2009).

La machinerie de méthylation de l'ADN semble exploiter les modifications locales de la chromatine pour réguler les profils de méthylation. En effet, de nombreuses études suggèrent que les Dnmts sont capables de "lire" les modifications des histones, entraînant leur recrutement au niveau de nucléosomes portant des marques spécifiques (Denis et al., 2011). Ainsi, grâce à leur domaine N-terminal, les Dnmt3s interagissent sélectivement avec l'histone



H3 non-méthylée sur la lysine 4 (H3K4me0), déclenchant une méthylation *de novo* au niveau des nucléosomes déficitaires en histone H3K4me (Hodges et al., 2009; Meissner et al., 2008; Ooi et al., 2007; Otani et al., 2009; Wu et al., 2010b; Zhang et al., 2010). D'autres modifications des histones, notamment H3K36me3 (Dhayalan et al., 2010; Hodges et al., 2009; Kolasinska-Zwierz et al., 2009; Laurent et al., 2010), H3K9me3 (Lehnertz et al., 2003) et H3K27me3 (Schlesinger et al., 2007; Vire et al., 2006), ont également été impliquées dans le guidage de la méthylation de l'ADN au niveau de régions spécifiques de la chromatine. Contrairement à H3K4, ces histones doivent être triméthylées, et certaines nécessitent la présence de protéines interagissant avec les Dnmt3a/3b telles que l'histone méthyltransférase EZH2 (H3K27me3) ou HP1 $\alpha$  (H3K9me3) pour induire le recrutement des Dnmt3s. En reconnaissant les modifications au niveau de la queue N-terminale des histones, les Dnmt3s peuvent ainsi être recrutées ou exclues de manière spécifique au niveau de nucléosomes portant des modifications distinctes (Chen and Riggs, 2011; Denis et al., 2011; Jones and Liang, 2009; Jurkowska et al., 2011; Smallwood and Kelsey, 2012).

Les méthyltransférases peuvent être recrutées au niveau de séquences spécifiques par d'autres protéines de liaison à l'ADN ou à la chromatine telles que des enzymes ou des protéines architecturales de la chromatine (CAPs) (Jones and Liang, 2009; Smallwood and Kelsey, 2012). Ainsi, il a été suggéré que les histones méthyltransférases (HMTs) EZH2 ("enhancer zeste homologue 2", un composant du complexe Polycomb PRC2) (Kondo et al., 2008; Vire et al., 2006) et G9a, recrutent Dnmt3a et Dnmt3b (Dong et al., 2008; Epsztejn-Litman et al., 2008). D'autre part, l'interaction directe des Dnmt3s avec des facteurs de transcription semble jouer un rôle important dans la méthylation ciblée de régions du génome qui expliquerait l'inhibition de l'expression de gènes induite par certains facteurs de transcription (Hervouet et al., 2009). Plusieurs autres protéines, telles que la protéine de l'hétérochromatine 1 (HP1) (Fuks et al., 2003a; Smallwood et al., 2007) et les désacétylases des histones (HDACs) (Robertson et al., 2000a) joueraient également un rôle dans le recrutement des Dnmts au niveau de régions particulières de la chromatine.

Particulièrement dans le cas de l'hétérochromatine, les Dnmts font face à un problème d'accessibilité de l'ADN qui se retrouve enroulé autour des histones. Le remodelage de la chromatine pourrait donc déterminer les régions du génome sujettes à la méthylation de l'ADN. Ainsi, les complexes de remodelage de la chromatine ATP-dépendants sont requis pour réguler l'accessibilité des Dnmts à leur substrat et pour permettre une méthylation

efficace de l'ADN précédemment condensé (Jurkowska et al., 2011; Smallwood and Kelsey, 2012).

De récentes études suggèrent que de petits ARNs pourraient servir de guide pour diriger la méthylation de l'ADN au niveau de sites génomiques spécifiques. Cette méthylation de l'ADN dirigée par l'ARN (RdDM, "RNA-directed DNA methylation") est observée chez les plantes, la levure et la drosophile, et implique l'interaction des DRM1/2 (homologues des Dnmt3a/3b) avec des ARNs non-codants (ARNnc) qui permettent de guider les enzymes au niveau de zones spécifiques du génome pour induire l'inhibition transcriptionnelle de gènes (TGS, "transcriptional gene silencing") par méthylation (Denis et al., 2011; Holz-Schietinger and Reich, 2012; Verdell et al., 2009). Les mécanismes de méthylation de l'ADN étant généralement similaires chez les plantes et les mammifères, l'hypothèse de l'existence d'une méthylation de l'ADN dirigée par l'ARN chez ces derniers a été proposée. Ainsi, de récentes études ont démontré qu'une classe de petits ARNnc, les ARNs interagissant avec Piwi (ARNpi), permettrait de diriger la méthylation de l'ADN au niveau de régions spécifiques du génome chez les mammifères. Ces ARNpi sont essentiellement exprimés dans les cellules germinales et sont caractérisés par leur interaction avec les protéines Piwi, une branche de la famille des protéines Argonaute. Dans la méthylation de l'ADN dirigée par l'ARN, ils interviendraient dans le recrutement des enzymes Dnmt3a et Dnmt3b pour induire la méthylation *de novo* au niveau des transposons (Aravin et al., 2008; Denis et al., 2011; Kuramochi-Miyagawa et al., 2008). D'autre part, des études *in vitro* ont permis d'observer la formation de complexes RNP ("RNA protein complexes") entre les Dnmt3a et Dnmt3b et l'ARN pouvant entraîner la modulation de l'activité catalytique des enzymes (Denis et al., 2011; Holz-Schietinger and Reich, 2012; Jeffery and Nakielny, 2004). D'autres études soutiennent l'hypothèse d'une méthylation de l'ADN médiée par l'ARN. Ainsi, une interaction entre de petits ARNnc et la Dnmt3a a été observée (Weinberg et al., 2006) et dans certains cas un petit ARNnc permet de diriger spécifiquement la méthylation de promoteurs par la Dnmt3b grâce à la formation d'une structure triplex ARN:ADN:ADN séquence-spécifique (Schmitz et al., 2010). Des études plus approfondies seront nécessaires pour élucider les mécanismes précis de méthylation de l'ADN dirigée par l'ARN chez les mammifères. Toutefois, ces études *in vitro*, combinées aux observations sur l'ARNpi, suggèrent que l'ARN joue un rôle important dans le recrutement des Dnmts au niveau de séquences spécifiques du génome et donc dans l'établissement des profils de méthylation de l'ADN (Denis et al., 2011).

Ainsi, afin d'établir les profils de méthylation de l'ADN au cours du développement, les Dnmt3s utilisent une multitude d'interactions (reconnaissance de séquences spécifiques de l'ADN, interactions avec les queues des histones modifiées, association avec des facteurs de remodelage de la chromatine, association à l'ARN ...) qui agissent séparément ou en synergie pour cibler des régions particulières du génome (Smallwood and Kelsey, 2012).

## 2.5. La méthylation de maintenance

La méthylation de maintenance permet la duplication des profils de méthylation de l'ADN lors de la réplication, assurant ainsi leur maintien pendant toute la durée de vie de l'organisme. C'est la Dnmt de maintenance, Dnmt1, qui assure l'essentiel de cette transmission. En effet, la Dnmt1, qui présente une affinité préférentielle pour l'ADN hémi-méthylé (Fatemi et al., 2001; Goyal et al., 2006; Hermann et al., 2004), est recrutée au niveau de la fourche de réplication via son interaction avec PCNA (Chuang et al., 1997; Easwaran et al., 2004), où elle copie le patron de méthylation présent sur le brin d'ADN "parental" afin de restaurer le profil de méthylation originel sur le brin d'ADN nouvellement synthétisé. Cependant, l'interaction de la Dnmt1 avec PCNA est transitoire (Egger et al., 2006; Schermelleh et al., 2007) et n'est pas indispensable pour son recrutement au niveau de l'ADN. En effet, la Dnmt1 est une enzyme hautement processive, et après son ancrage à l'ADN, elle peut méthyliser de grandes étendues d'ADN sans dissociation, par diffusion linéaire (Goyal et al., 2006; Hermann et al., 2004; Vilkaitis et al., 2005). Un autre partenaire important dans le recrutement de la Dnmt 1 au niveau de l'ADN est la protéine UHRF1 ("Ubiquitin-like plant Homeodomain and Ring Finger 1"). Cette protéine se lie à l'ADN hémi-méthylé (Arita et al., 2008; Avvakumov et al., 2008; Bostick et al., 2007; Hashimoto et al., 2008; Sharif et al., 2007) et colocalise avec PCNA et la Dnmt1 au niveau de la fourche de réplication (Bostick et al., 2007; Sharif et al., 2007; Uemura et al., 2000). En interagissant directement avec la Dnmt1, UHRF1 aiderait au recrutement de l'enzyme au niveau de la fourche de réplication. Le complexe Dnmt1/PCNA/UHRF1 est principalement formé au cours de la phase S du cycle cellulaire et permet la maintenance de la méthylation lors de la réplication de l'ADN (Hervouet et al., 2012). Cependant, la duplication des profils de méthylation de l'ADN peut également avoir lieu lors des phases G0/G1 et G2/M du cycle cellulaire grâce à des facteurs de transcriptions qui permettraient le recrutement de la Dnmt1 au niveau de sites spécifiques

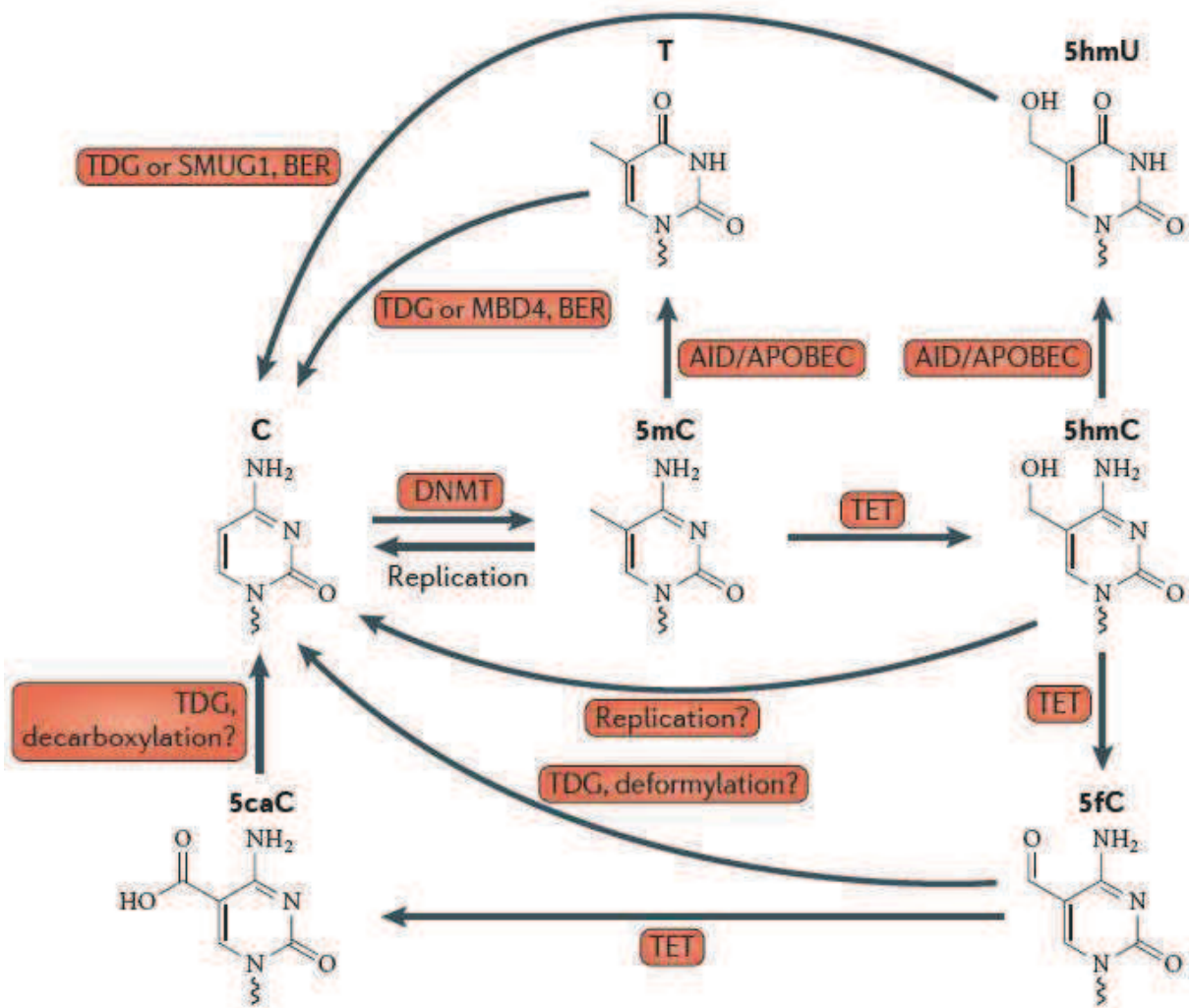
du génome (Hervouet et al., 2012; Hervouet et al., 2010b). D'autre part, la protéine UHRF1 permettrait de diriger la Dnmt1 au niveau des sites CpG oubliés pendant la réplication afin de corriger ses erreurs et de compléter la méthylation du brin d'ADN nouvellement synthétisé (Jones and Liang, 2009; Jurkowska et al., 2011).

Bien que la Dnmt1 soit principalement responsable de la méthylation de maintenance, ce processus requiert également la coopération des enzymes *de novo* Dnmt3a et Dnmt3b (Arand et al., 2012; Chen et al., 2003; Dodge et al., 2005; Jeong et al., 2009; Kim et al., 2002; Liang et al., 2002). Dans les cellules somatiques, la Dnmt3a et la Dnmt3b ne semblent pas exister sous forme soluble dans le noyau mais sont ancrées de manière stable aux nucléosomes contenant une forte concentration d'ADN méthylé (Jeong et al., 2009). Ainsi, la Dnmt3a et la Dnmt3b resteraient associées à des régions spécifiques du génome grâce à des interactions au niveau de nucléosomes portant des modifications distinctes. Après le passage de la fourche de réplication, ces deux enzymes permettraient de méthyler les sites oubliés par la Dnmt1 (Jones and Liang, 2009). D'autre part, il a été démontré que la protéine UHRF1 était capable de diriger la Dnmt3a et la Dnmt3b au niveau des sites d'ADN hémi-méthylés (Meilinger et al., 2009), cette interaction pouvant également contribuer à la méthylation de maintenance par les Dnmt3s.

Ainsi, la transmission des profils de méthylation est assurée par la combinaison des mécanismes d'action de la Dnmt de maintenance et des *de novo* Dnmt3a et Dnmt3b. La Dnmt3a et la Dnmt3b semblent compartimentées au niveau des régions richement méthylées telles que les îlots CpG et les séquences répétées, alors que la Dnmt1 assure la majeure partie de la méthylation de l'ADN localisée en dehors des îlots CpG (Jones and Liang, 2009).

## 2.6. La déméthylation de l'ADN

Historiquement, la méthylation de l'ADN est considérée comme une modification relativement stable puisqu'elle est transmise de cellule en cellule durant toute la durée de vie de l'organisme. Toutefois, la reprogrammation épigénétique, qui induit la perte et la modification des profils de méthylation après la fécondation et dans les cellules germinales primordiales, révèle l'existence de mécanismes de déméthylation de l'ADN et témoigne clairement de la nature dynamique de la méthylation de l'ADN.



*Figure 10 : Les voies de déméthylation potentielles (Branco et al., 2011). La 5-méthylcytosine (5mC) peut être éliminée de manière passive pendant la réplication, mais plusieurs voies de déméthylation actives ont aussi été proposées, incluant parfois la 5-hydroxyméthylcytosine (5hmC) comme intermédiaire. La désamination de la 5mC ou la 5hmC via l'action de la famille des désaminases de la cytidine AID/APOBEC produit des mésappariements reconnus par les glycosylases de l'ADN qui forment des sites abasiques corrigés grâce au système de réparation par excision de base (BER). Alternativement, la 5hmC peut être davantage oxydée en 5-formylcytosine (5fC) et en 5-carboxylcytosine (5cac) par les enzymes TET. Bien qu'il soit possible qu'une déformylation de la 5fC et une décarboxylation de la 5caC reconvertisent directement ces intermédiaires en cytosine, aucune activité enzymatique de ce type n'a encore été découverte. En revanche, il a été démontré que la glycosylase de la thymine de l'ADN (TDG) est capable de cliver la 5fC et la 5caC, impliquant à nouveau l'action de la voie BER dans la déméthylation de l'ADN. 5hmU, 5-hydroxyméthyluracile; DNMT, méthyltransférase de l'ADN; MBD4, "methyl-CpG-binding domain protein 4"; SMUG1, "single-strand-selective monofunctional uracil-DNA glycosylase 1".*

Les mécanismes de déméthylation de l'ADN ont été soumis à controverse (Ooi and Bestor, 2008), toutefois, de récentes études suggèrent qu'elle s'effectue selon deux processus: passif ou actif. La déméthylation passive se produit par un déficit de la méthylation de maintenance lors de la réplication, induisant la perte progressive des méthylations à chaque division cellulaire. A l'inverse, la déméthylation active dépend de l'activité d'une ou plusieurs enzymes et peut se produire indépendamment de la réplication de l'ADN (Bhutani et al., 2011; Branco et al., 2011; You and Jones, 2012). Ainsi, plusieurs études ont permis de mettre en évidence l'existence de trois familles d'enzymes qui agissent en synergie pour induire la déméthylation active de l'ADN (Figure 10): la famille Tet ("ten-eleven-translocation"), qui modifie les 5-méthylcytosines (5mC) d'abord par hydroxylation puis par des oxydations successives; la famille AID/APOBEC, qui désamine les bases (5mC ou 5hmC); et des glycosylases qui activent le système de réparation par excision de base (BER, "base excision repair") (Bhutani et al., 2011; Branco et al., 2011).

Chez les plantes, les glycosylases de l'ADN peuvent éliminer la 5mC, créant un site abasique qui peut ensuite être corrigé par le système BER. En revanche, chez les mammifères, les glycosylases de l'ADN connues, telles que la glycosylase de la thymine de l'ADN (TDG, "thymine DNA glycosylase") et la protéine MBD4, n'exercent pas directement leur activité catalytique sur la 5mC. Ces enzymes montrent une forte activité envers les mésappariements T:G créés à la suite de la désamination des 5mC par la famille des désaminases de la cytidine AID/APOBEC. Ainsi, chez les mammifères, la déméthylation de la 5mC aurait lieu en deux étapes, une déamination intermédiaire précédant l'action des glycosylases et la réparation des mésappariements par le système BER (Bhutani et al., 2011; Branco et al., 2011; Cortellino et al., 2011; Guo et al., 2011).

La 5-hydroxyméthylcytosine (5hmC), qui provient de l'oxydation du groupement méthyle de la 5-méthylcytosine (5mC), a été récemment découverte dans le génome de mammifères (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009). C'est la famille de protéines Tet (Tet1 à Tet3) qui est responsable de la conversion catalytique de la 5mC en 5hmC, des oxydations itératives conduisant à la formation d'autres dérivés de l'oxydation, tels que la 5-formylcytosine (5fC) et la 5-carboxylcytosine (5caC) (Bhutani et al., 2011; Branco et al., 2011; Dawson and Kouzarides, 2012; He et al., 2011; Ito et al., 2011; Pfaffeneder et al., 2011). La 5fC et la 5caC peuvent toutes deux être réparées par TDG, impliquant à nouveau l'induction du système BER dans cette voie de déméthylation de l'ADN (Bhutani et al., 2011; Branco et al., 2011; He et al., 2011; Maiti and Drohat, 2011). D'autre part, on suppose



l'existence d'enzymes déformylase et décarboxylase qui pourraient directement convertir la 5fC et la 5caC en cytosine, indépendamment du système BER (Ito et al., 2011). Le complexe enzymatique des désaminases AID/APOBEC peut également agir sur la 5hmC, générant des mésappariements U:G pouvant être réparés par l'action consécutive des glycosylases SMUG1 ("single-strand-selective monofunctional uracil-DNA glycosylase 1") et TDG et du système BER (Branco et al., 2011; Cortellino et al., 2011; Guo et al., 2011; He et al., 2011).

Récemment, il a été démontré que la 5hmC pouvait être perdue de manière passive lors des divisions cellulaires (Inoue and Zhang, 2011). En effet, la conversion de la 5mC en 5hmC pourrait favoriser la déméthylation passive en empêchant l'action de la Dnmt1 sur les sites CpG hydroxyméthylés. Cependant, bien que la Dnmt1 ne soit pas active *in vitro* sur les sites CpG hémi-hydroxyméthylés, sa protéine partenaire UHRF1 se lie avec la même affinité à la 5mC et la 5hmC (Frauer et al., 2011) ce qui pourrait favoriser l'activité de la Dnmt1 au niveau des sites CpG hydroxyméthylés (Branco et al., 2011). Des études plus approfondies seront nécessaires pour clarifier la fonction de UHRF1 dans ces mécanismes de déméthylation.

Ainsi, bien que leur rôle reste encore à éclaircir, les dérivés d'oxydation de la 5mC sont des intermédiaires essentiels au processus de déméthylation actif et passif de l'ADN. D'autre part, la 5hmC pourrait être une modification épigénétique à part entière. En effet, elle possède une distribution distincte de la 5mC dans le génome au niveau de certains promoteurs, et module l'expression des gènes en altérant le recrutement de régulateurs de la chromatine tels que les protéines MBPs ou encore des complexes répresseurs Polycomb (Branco et al., 2011 ; Wu et al., 2011; Wu and Zhang, 2011).

## 2.7. Rôles physiologiques de la méthylation de l'ADN

La méthylation de l'ADN est un mécanisme épigénétique clé qui intervient dans de nombreux processus physiologiques essentiels tels que le développement et la différenciation cellulaire, l'empreinte génomique parentale, l'inactivation du chromosome X, et l'adaptation à l'environnement.

Au cours du développement, le zygote et les cellules germinales primordiales subissent une reprogrammation épigénétique afin d'acquérir de nouveaux profils de

méthylation spécifiques responsables de la différenciation cellulaire. Ainsi, lorsque les cellules évoluent du stade de totipotence en passant par des étapes de pluripotence pour aboutir à un état bien différencié, on observe un changement progressif des profils de méthylation permettant de réduire au silence les gènes des progéniteurs au profit de l'expression des gènes tissu-spécifiques. Des études génomiques de ces profils de méthylation ont ainsi révélé qu'une faible proportion d'îlots CpG peuvent adopter un état de méthylation tissu-spécifique (Eckhardt et al., 2006; Futscher et al., 2002; Illingworth et al., 2008; Imamura et al., 2001; Irizarry et al., 2009; Lister et al., 2009; Maegawa et al., 2010; Meissner et al., 2008; Schilling and Rehli, 2007; Shen et al., 2007; Shiota et al., 2002; Song et al., 2005; Straussman et al., 2009; Zhang et al., 2009) (Illingworth and Bird, 2009; Jurkowska et al., 2011). En effet, ce sont surtout les séquences peu denses en CpG tels que les "CpG island shores", contenant des régions de méthylation différentielle (DMR), qui sont impliquées dans la différenciation cellulaire (Fernandez et al., 2012a; Hodges et al., 2011; Irizarry et al., 2009; Wu et al., 2010a). La méthylation de l'ADN représente un mécanisme important de régulation tissu-spécifique car elle constitue une marque épigénétique stable, hautement transmissible, associée à la répression transcriptionnelle, et qui contribue à l'établissement de lignées différenciées stables.

Malgré la reprogrammation épigénétique observée lors des premiers stades du développement embryonnaire, certaines régions du génome impliquées dans l'établissement de l'empreinte génomique parentale échappent à la vague de déméthylation (Koukoura et al., 2012; Mayer et al., 2000; Smallwood and Kelsey, 2012; Tremblay et al., 1995). Dans les cellules diploïdes de mammifères, la plupart des gènes autosomiques sont exprimés équitablement à partir des allèles paternel et maternel, entraînant une expression biallélique de ces gènes. Cependant, une faible proportion de gènes est soumise à l'empreinte génomique parentale, un processus par lequel un allèle parental est réduit au silence dans le zygote, conduisant à une expression monoallélique (Ishida and Moore, 2012). Lors de l'établissement de ces empreintes, la lignée germinale mâle ou femelle confère des marques spécifiques du genre au niveau de certaines régions du chromosome (Koukoura et al., 2012; Reik and Walter, 2001). Seul l'un des deux allèles des gènes soumis à empreinte, l'allèle maternel ou paternel, peut être actif et exprimé. Actuellement, on compte approximativement 60 gènes soumis à empreinte chez l'homme (Glaser et al., 2006). Typiquement, les gènes soumis à empreinte sont organisés en pôles et contrôlés par des régions de contrôle d'empreinte ICR ("imprinting control region") qui agissent en *cis* pour réguler l'expression des gènes (Fang et al., 2012;



Ferguson-Smith, 2011). Ces domaines ICR contiennent des régions de méthylation différentielle (DMR) dont les profils de méthylation spécifiquement établis pendant la gamétogenèse (DMR germinale) ou après la fécondation (DMR somatique) sont essentiels au maintien de l'expression monoallélique des gènes soumis à empreinte (Davis et al., 2000; Edwards and Ferguson-Smith, 2007; El-Maarri et al., 2001; Fang et al., 2012; Ferguson-Smith, 2011; Ishida and Moore, 2012; Koukoura et al., 2012; Lewis and Reik, 2006; Wood and Oakey, 2006).

La méthylation de l'ADN intervient dans des étapes précises du développement comme l'inactivation du chromosome X. Ce processus, qui se produit chez les mammifères femelles (XX), permet l'inactivation transcriptionnelle d'un des deux chromosomes X afin d'égaliser le niveau d'expression des gènes de l'unique chromosome X présent chez les mâles (XY). Ce phénomène constitue un événement épigénétique remarquable dans lequel deux chromosomes parfaitement identiques en termes de séquence d'ADN se distinguent uniquement par leurs modifications épigénétiques. Ainsi, des marques épigénétiques telles que des modifications des histones, des variants d'histones, la méthylation de l'ADN, sont toutes enrichies sur l'un des deux chromosomes X, le chromosome X inactivé (Xi), afin de réduire ses gènes au silence (Basu and Zhang, 2011). L'inactivation du chromosome Xi est principalement régulée via le gène Xist ("X-inactive specific transcript"), qui transcrit un long ARNnc exprimé à partir du chromosome Xi pendant le développement embryonnaire précoce puis de façon aléatoire dans les cellules somatiques femelles. L'ARN Xist s'associe en *cis* sur le chromosome Xi en le recouvrant et permet le recrutement d'autres facteurs de répression génique afin d'établir une répression spécifique du chromosome X (Pontier and Gribnau, 2011).

De nombreuses études tendent à montrer que les facteurs environnementaux peuvent moduler l'établissement et la maintenance des modifications épigénétiques sans mutation génétique sous-jacente. Ainsi, les polluants chimiques, le régime alimentaire, les changements de température, et d'autres stress externes peuvent avoir un effet durable sur le développement, le métabolisme et la santé, parfois même sur plusieurs générations (Feil and Fraga, 2012). L'impact des facteurs extrinsèques sur l'épigénome des mammifères semble dépendre du stade de développement et paraît particulièrement important lors de la gestation, l'environnement ayant un effet sur le placenta et l'embryon. Ainsi, le régime alimentaire de la mère ou son exposition à des composés toxiques durant le développement embryonnaire est associé à des changements épigénétiques au niveau de régions spécifiques du génome avec

l'apparition de profils de méthylation particuliers (Daxinger and Whitelaw, 2012; Feil and Fraga, 2012; Gluckman et al., 2009; Heijmans et al., 2008; Rosenfeld, 2010; Tobi et al., 2009; Waterland et al., 2010). Ces modifications épigénétiques sont parfois transmissibles et peuvent avoir un impact sur la santé de la descendance (Daxinger and Whitelaw, 2012; Feil and Fraga, 2012; Heijmans et al., 2008; Kaati et al., 2002; Lumey et al., 2009; Pembrey et al., 2006). Des facteurs environnementaux peuvent également affecter l'épigénome de manière post-natal. Ainsi, des études épidémiologiques ont révélé des changements dans les profils de méthylation d'individus ayant été exposés à la pollution de l'air ou au benzène, à la fumée de tabac, à une exposition chronique au soleil, ou à une forte consommation d'alcool (Baccarelli et al., 2009; Bollati et al., 2007; Christensen et al., 2009; Feil and Fraga, 2012; Gronniger et al., 2010; Langevin et al., 2011). Dans certains cas, les marques épigénétiques acquises pendant le développement d'un individu peuvent être transmises sur plusieurs générations. On observe ainsi une centaine de gènes qui ne sont pas soumis à empreinte dont la méthylation au niveau de leur promoteur est constante et échappe à la reprogrammation épigénétique (Borgel et al., 2010). D'autre part, certaines classes de rétrotransposons, en particulier les éléments IAPs ("intracisternal A-type particles") sont également résistantes à la reprogrammation épigénétique (Hajkova et al., 2008; Hajkova et al., 2002; Lane et al., 2003; Popp et al., 2010). L'existence de certaines régions du génome, échappant à la déméthylation lors du développement, suggère l'implication de la méthylation de l'ADN dans la transmission de l'information épigénétique à la descendance via les gamètes (Daxinger and Whitelaw, 2012; Feil and Fraga, 2012).

## 2.8. Dérégulation de la méthylation de l'ADN

De nombreuses études ont révélé un lien entre des altérations du système de méthylation de l'ADN et des maladies neurologiques. Ainsi, des mutations dans le gène codant pour la protéine MeCP2 sont responsables d'un désordre postnatal neuro-développemental lié au chromosome X, le syndrome de Rett. Cette maladie touche essentiellement les filles et se caractérise, entre autres, par un retard mental sévère et des anomalies motrices. D'autres retards mentaux, incluant le syndrome ATRX ("alpha thalassaemia/mental retardation X-linked") et le syndrome de l'X fragile, sont associés à une hyper- ou une hypo-méthylation du promoteur de gènes spécifiques entraînant leur inhibition

ou leur surexpression. Des profils de méthylation aberrants sont également associés à la maladie d'Alzheimer, de Parkinson, l'ataxie de Friedreich, l'amyotrophie spinale, et d'autres désordres neurologiques tels que la sclérose en plaques (Portela and Esteller, 2010; Urdinguio et al., 2009). Outre l'altération des profils de méthylation au niveau de promoteurs de gènes, des maladies neurologiques peuvent également être associées à une perte d'empreinte génomique parentale (Ishida and Moore, 2012; Sasai and Defossez, 2009). Les syndromes de Prader-Willi et d'Angelman en sont un bon exemple, le syndrome de Prader-Willi étant associé à la perte de l'expression des gènes paternels, et le syndrome d'Angelman, à la perte d'expression des gènes maternels d'une même région soumise à empreinte (Portela and Esteller, 2010).

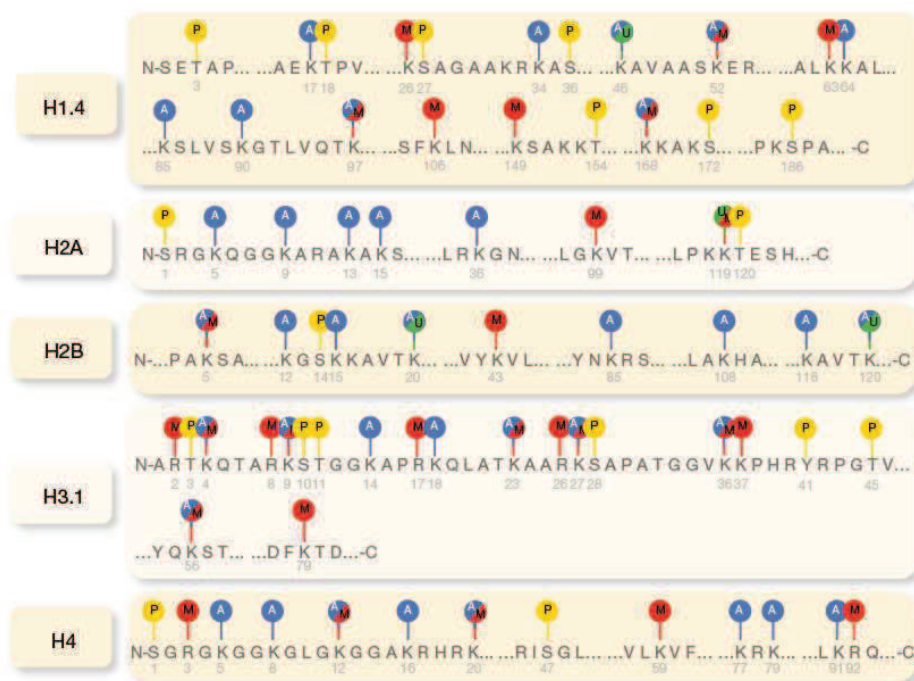
Des altérations dans la méthylation de l'ADN ont également été observées dans des maladies auto-immunes, la mieux connue étant le syndrome ICF ("immunodeficiency, centromeric instability and facial anomalies") généré par des mutations du gène codant pour la Dnmt3b (Jones, 2012; Portela and Esteller, 2010; Sasai and Defossez, 2009). Les patients atteints d'ICF présentent une hypométhylation des séquences répétitives péri-centromériques qui entraîne une instabilité de l'hétérochromatine centromérique. Cette maladie génétique est également caractérisée par une immunodéficience et des déformations faciales. D'autres maladies auto-immunes comme le lupus érythémateux disséminé et la polyarthrite rhumatoïde sont associées à une hypométhylation globale du génome (Portela and Esteller, 2010).

L'établissement de profils aberrants de méthylation de l'ADN ainsi que des mutations touchant les protéines de reconnaissance de l'ADN méthylé sont donc impliqués dans diverses maladies. Dans le cancer, outre les mutations génétiques classiques caractéristiques, les cellules cancéreuses présentent de profondes variations épigénétiques qui seront détaillées plus loin (voir p. 50).

### 3. Les modifications post-traductionnelles des histones

Grâce à leur domaine C-terminal globulaire, les protéines histones forment le cœur des nucléosomes, l'unité de base de la chromatine. Leur extrémité N-terminale émergeant de la surface des nucléosomes est la cible privilégiée de nombreuses modifications covalentes post-traductionnelles telles que l'acétylation, la méthylation, la phosphorylation, l'ubiquitinylation

ou encore la sumoylation (Figure 11) (Bannister and Kouzarides, 2011; Kouzarides, 2007; Portela and Esteller, 2010; Rando and Chang, 2009; Sharma et al., 2010; Tollervey and Lunyak, 2012). Ces modifications constituent un processus clé dans la régulation épigénétique puisqu'elles vont permettre de moduler l'expression des gènes en influençant directement l'accessibilité de la chromatine ou en créant des sites de liaison spécifiques pour d'autres protéines ou des complexes enzymatiques (Bannister and Kouzarides, 2011; Sharma et al., 2010). Ainsi, selon le résidu concerné et le type de modification présente, les modifications des histones peuvent conduire à l'activation ou la répression de gènes. En général, l'acétylation des lysines entraîne l'activation de la transcription alors que leur méthylation induit l'activation ou la répression des gènes selon le résidu modifié et le degré de méthylation (mono-, di- ou tri-méthylé) (Kouzarides, 2007; Sharma et al., 2010). Les modifications des histones ont lieu au niveau de régions spécifiques du génome, générant ainsi des plages dépourvues ou enrichies de modifications caractéristiques. Ainsi, l'euchromatine est caractérisée par un niveau élevé d'acétylation et de tri-méthylation sur H3K4, H3K36 et H3K79 alors que l'hétérochromatine est hypoacétylée mais hyperméthylée sur les sites H3K9, H3K27 et H4K20 (Li et al., 2007; Portela and Esteller, 2010).



*Figure 11 : Modifications post-traductionnelles des histones (Portela and Esteller, 2010). Toutes les histones sont sujettes à des modifications post-traductionnelles qui ont principalement lieu au niveau de leurs queues N-terminales. Les principales modifications sont représentées sur cette figure: acétylation (bleu), méthylation (rouge), phosphorylation (jaune) et ubiquitinylation (vert). Le nombre en gris situé en-dessous de chaque acide aminé représente leurs positions dans la séquence.*

Les modifications des histones jouent donc un rôle important dans la régulation de la transcription. D'autre part, puisqu'elles affectent l'état de la chromatine, elles peuvent également influencer d'autres processus tels que la réparation de l'ADN (Huertas et al., 2009), la réplication, la recombinaison et l'épissage alternatif (Bannister and Kouzarides, 2011; Luco et al., 2010; Portela and Esteller, 2010; Sharma et al., 2010).

### 3.1. L'acétylation des histones

L'acétylation des histones est un processus dynamique qui se produit sur les résidus lysines. Elle est gouvernée par deux familles d'enzymes antagonistes, les histones acétyltransférases (HATs) et les histones déacétylases (HDACs) (Bannister and Kouzarides, 2011; Lee and Workman, 2007).

Les HATs catalysent le transfert d'un groupement acétyl du cofacteur acétyl-CoA vers le groupement  $\epsilon$ -aminé de la chaîne latérale de résidus lysines spécifiques. Il en résulte une neutralisation de la charge positive des résidus lysines qui affaiblit l'interaction électrostatique entre l'ADN et les histones cibles, la chromatine adoptant alors une structure relaxée plus permissive à la transcription (Bannister and Kouzarides, 2011; Yang and Seto, 2007). Ainsi, parmi les divers sites spécifiques, l'acétylation de la lysine 16 de l'histone H4 (H4K16) apparaît comme un événement essentiel dans la régulation de l'état de compaction de la chromatine et dans l'établissement et la propagation de l'hétérochromatine (Arrowsmith et al., 2012; Shahbazian and Grunstein, 2007; Woodcock and Ghosh, 2010). Outre son action directe sur l'accessibilité de la chromatine, l'acétylation des histones favorise également la transcription en permettant la création de sites de liaison pour des protéines impliquées dans l'activation des gènes. En particulier, la famille des protéines à bromodomaine s'associe aux HATs pour former des complexes de remodelage de la chromatine au niveau des lysines acétylées (Arrowsmith et al., 2012; Bannister and Kouzarides, 2011; Yang and Seto, 2007). La nature des protéines composant ces gros complexes multiprotéiques influence le recrutement de l'enzyme HAT, son activité et sa spécificité (Bannister and Kouzarides, 2011).

A l'inverse des HATs, les HDACs annulent l'acétylation des résidus lysines en restaurant leur charge positive et en stabilisant l'architecture de la chromatine. Ainsi, les HDACs sont principalement associées à des complexes répresseurs. Contrairement aux

HATs, ces enzymes sont peu spécifiques et peuvent déacétyler de multiples sites au sein des histones (Arrowsmith et al., 2012; Bannister and Kouzarides, 2011).

### 3.2. La méthylation des histones

La méthylation des histones se produit sur les résidus lysines qui peuvent être mono-, di- ou tri-méthylés, ou sur les résidus arginines qui peuvent être mono- ou di-méthylés (symétriquement ou asymétriquement) (Bannister and Kouzarides, 2011; Bedford and Clarke, 2009; Ng et al., 2009). Comme l'acétylation, la méthylation des histones est un processus dynamique qui requiert deux types d'enzymes antagonistes. Les histones méthyltransférases (HMTs), grâce à leur domaine SET, catalysent le transfert d'un groupement méthyl du cofacteur SAM vers le groupement  $\epsilon$ -aminé de la chaîne latérale de résidus lysines (PKMTs, "protein lysine methyltransferase") ou le groupement  $\omega$ -guanidine de résidus arginines (PRMTs, "protein arginine methyltransferase"). A l'inverse, les groupements méthyls sont éliminés par deux familles d'enzymes, la famille Jumonji des déméthylases dépendantes du 2-oxoglutarate (Tsukada et al., 2006) et les déméthylases des lysines dépendantes des flavines (LSD, "lysine-specific histone demethylase") (Arrowsmith et al., 2012; Shi et al., 2004).

La méthylation des histones a lieu sur des résidus caractéristiques selon la spécificité de l'enzyme HMT qui est sensible au degré de méthylation préalable du résidu cible et le modifie d'un degré qui lui est propre (mono-, di- ou tri-méthylation) (Bannister and Kouzarides, 2011). Contrairement à l'acétylation, la méthylation des histones n'altère pas la charge de la protéine histone et n'a donc pas d'effet direct sur la structure de la chromatine (Arrowsmith et al., 2012; Bannister and Kouzarides, 2011). Ainsi, en constituant des sites de fixation pour des protéines et des complexes spécifiques, la méthylation des histones entraîne selon le cas l'activation ou la répression des gènes. Par exemple, la tri-méthylation de la lysine 4 sur l'histone H3 (H3K4me3) est associée aux promoteurs transcriptionnellement actifs (Liang et al., 2004) alors que la triméthylation d'H3K9 (H3K9me3) et H3K27 (H3K27me3) est présente au niveau du promoteur de gènes réprimés (Arrowsmith et al., 2012; Bannister and Kouzarides, 2011; Kouzarides, 2007; Sharma et al., 2010).



### 3.3. Autres modifications post-traductionnelles des histones

Outre l'acétylation et la méthylation, les histones sont sujettes à de multiples autres modifications telles que la phosphorylation, l'ubiquitinylation, la sumoylation, l'ADP-ribosylation ou la biotinylation, et de nouveaux types de modifications post-traductionnelles sont découverts régulièrement. Seules certaines de ces modifications seront décrites dans ce manuscrit.

La phosphorylation des histones est un processus dynamique gouverné par les kinases et les phosphatases qui a principalement lieu sur les résidus sérines, thréonines et tyrosines (Oki et al., 2007). De manière identique à l'acétylation, la phosphorylation influence directement l'architecture de la chromatine en modifiant la charge des histones (Bannister and Kouzarides, 2011). D'autre part, elle semble influencer les autres modifications post-traductionnelles en stimulant ou en inhibant l'activité des enzymes des histones.

L'ubiquitinylation est une modification covalente considérable puisqu'elle permet l'ajout d'un ou plusieurs polypeptides de 76 acides aminés aux lysines via l'action consécutive de l'enzyme d'activation E1, de conjugaison E2 et de ligation E3. Malgré la taille importante de l'ubiquitinylation, ce processus est réversible grâce à l'action des isopeptidases. Alors que la poly-ubiquitinylation destine les protéines à la dégradation par le protéasome, la mono-ubiquitinylation des histones est stable et participe à des fonctions régulatrices. Considérant la taille de la modification, l'ubiquitinylation induit une restructuration des nucléosomes et de la chromatine qui perturbe les interactions de ses partenaires. Elle peut ainsi influencer l'expression des gènes (Wang et al., 2004) ou jouer un rôle dans l'initiation de la transcription et l'élongation (Bannister and Kouzarides, 2011; Kim et al., 2009a; Lee et al., 2007).

La sumoylation est une modification apparentée à l'ubiquitinylation qui implique la liaison covalente de petites molécules SUMO ("small ubiquitin-like modifier") via l'action des enzymes E1, E2 et E3. Bien que ses mécanismes d'action restent à élucider, la sumoylation semble associée à la répression transcriptionnelle (Shiio and Eisenman, 2003).

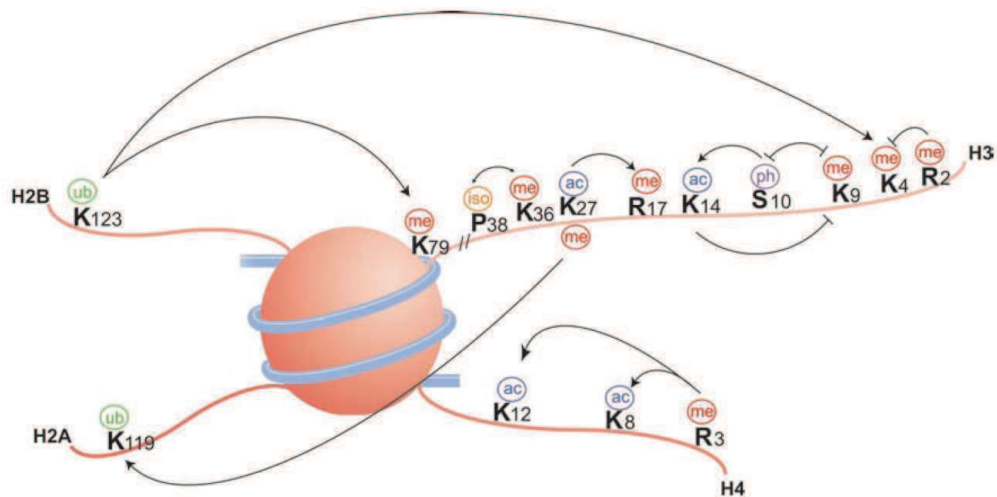
La mono- et la poly-ADP-ribosylation constituent des modifications dynamiques des résidus glutamate et arginine gouvernées par l'action combinée de la mono-ADP-ribosyltransférase, des poly-ADP-ribose polymérase (PARP) et des poly-ADP-ribose-



glycohydrolases. Ces modifications sont corrélées à un état relaxé de la chromatine et semblent jouer un rôle dans la réponse aux dommages de l'ADN (Hassa et al., 2006).

### 3.4. Le code histone

Les histones subissent des modifications post-traductionnelles sur divers sites simultanément. Collectivement, ces modifications constituent un "code histone" permettant de définir la structure et la fonction d'une région de la chromatine (Ernst and Kellis, 2010; Strahl and Allis, 2000). En effet, au sein du cœur d'histones formant le nucléosome, chaque modification peut induire ou inhiber d'autres modifications post-traductionnelles (Figure 12). Cette communication entre les modifications des histones peut avoir lieu en *cis*, c'est-à-dire sur la même histone (Duan et al., 2008; Wang et al., 2008), ou en *trans*, lorsqu'elle entraîne une modification sur un résidu d'une autre protéine histone (Nakanishi et al., 2009; Portela and Esteller, 2010). D'autre part, cette collaboration peut fonctionner sur un même nucléosome ou s'établir entre deux nucléosomes différents (Lee et al., 2010). La communication entre les diverses modifications des histones peut se produire via de multiples mécanismes (Figure 12): des modifications ciblant le même résidu peuvent entraîner un antagonisme compétitif (par exemple l'acétylation et la méthylation sont des modifications mutuellement exclusives); une modification peut être sous la dépendance de la modification préalable d'un résidu spécifique (par exemple l'ubiquitinylation d'H2B entraîne la méthylation d'H3K4) (Kim et al., 2009a); la liaison d'une protéine sur un résidu modifié particulier peut être perturbée par une modification adjacente (par exemple la liaison d'HP1 à H3K9 méthylé est inhibée par la phosphorylation d'H3S10) (Fischle et al., 2005); l'activité d'une enzyme peut être affectée par la modification préalable de son substrat (par exemple, chez les levures, l'isomérisation d'H3P38 empêche la méthylation d'H3K36 par scSET2) (Kiefer et al., 2008); plusieurs modifications peuvent coopérer afin de recruter des facteurs spécifiques qui les reconnaissent simultanément (par exemple PHF8 se lie préférentiellement à H3K4me3 lorsque H3K9 et H3K14 sont acétylés) (Bannister and Kouzarides, 2011; Vermeulen et al., 2010).



*Figure 12 : Inter-relations entre les modifications des histones (Bannister and Kouzarides, 2011). Les modifications des histones peuvent affecter positivement ou négativement d'autres modifications. Un effet positif est indiqué par une tête de flèche et un effet négatif par une tête plate.*

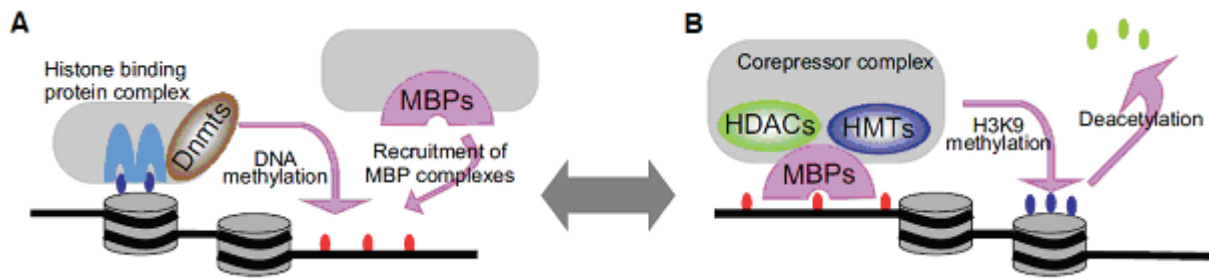
Ainsi, les modifications post-traductionnelles individuelles ou combinées des histones constituent des patrons d'information interprétés par de larges complexes multiprotéiques qui contribuent au remodelage de la chromatine et à l'établissement de la transcription (Arrowsmith et al., 2012; Sharma et al., 2010). L'influence des patrons de modifications d'histones sur l'état de la chromatine suggère des rôles biologiques distincts (Ernst and Kellis, 2010). En effet, des profils de modifications des histones sont spécifiquement présents dans certains types de cellules et joueraient un rôle dans le maintien de l'identité cellulaire (Mikkelsen et al., 2007). Ainsi, les cellules embryonnaires sont caractérisées par la présence simultanée de marques de répression H3K27me3 et d'activation H3K4me3 qui constituent des domaines bivalents (Sharma et al., 2010; Tollervey and Lunyak, 2012). Ces domaines particuliers de la chromatine sont perdus lors de la différenciation cellulaire et auraient une importance dans la régulation des gènes liés au développement (Bernstein et al., 2006; Mikkelsen et al., 2007).

### 3.5. Connexion entre la méthylation de l'ADN et les modifications des histones

Tous les protagonistes de la régulation épigénétique semblent interagir mutuellement. Ainsi, en plus d'accomplir leur rôles individuels, les acteurs des modifications des histones et de la méthylation de l'ADN communiquent à divers niveaux afin de réguler l'expression des gènes, l'organisation de la chromatine et l'identité cellulaire (Cedar and Bergman, 2009; Portela and Esteller, 2010; Sharma et al., 2010).

Comme nous l'avons vu précédemment, les modifications des histones seraient capables de diriger la méthylation de l'ADN au niveau de régions spécifiques du génome (Figure 13A). Ainsi, le statut de méthylation d'H3K4 jouerait un rôle important dans l'établissement de la méthylation de l'ADN puisque l'interaction spécifique des Dnmt3s au niveau de la queue de l'histone H3 est inhibée par la méthylation d'H3K4 (Hodges et al., 2009; Meissner et al., 2008; Ooi et al., 2007; Otani et al., 2009; Wu et al., 2010b; Zhang et al., 2010). D'autre part, en recrutant les Dnmts, plusieurs histones méthyltransférases (HMTs) telles que G9a, Suv39h1 et PRMT5 permettent de diriger la méthylation de l'ADN au niveau de régions spécifiques du génome, favorisant ainsi la répression des gènes (Dong et al., 2008; Epsztejn-Litman et al., 2008; Lehnertz et al., 2003; Tachibana et al., 2008; Zhao et al., 2009). Les histones méthyltransférases et déméthylases jouent davantage un rôle dans le contrôle de la méthylation de l'ADN en régulant la stabilité des Dnmts (Esteve et al., 2009; Wang et al., 2009).

A l'inverse, la méthylation de l'ADN peut influencer les modifications des histones (Figure 13B). Ainsi, les protéines de reconnaissance de l'ADN méthylé (MBPs) tel que MeCP2 recrute des histones déacétylases (HDACs) au niveau de l'ADN méthylé, favorisant la répression transcriptionnelle et la condensation de la chromatine (Jones et al., 1998; Nan et al., 1998). En plus de sa liaison à des HDACs, MeCP2 est impliqué dans la méthylation d'H3K9, stabilisant davantage la répression de la chromatine (Fuks et al., 2003b). La méthylation de l'ADN peut également empêcher l'action des enzymes de modification des histones. Ainsi, le recrutement de KDM2A, l'histone déméthylase spécifique d'H3K36, au niveau d'H3K9me3 est bloqué lorsque l'ADN est méthylé (Bartke et al., 2010).



*Figure 13 : Connexion entre la méthylation de l'ADN et les modifications des histones (Sasai and Defossez, 2009). (A) Modèle de méthylation de l'ADN dirigée par les modifications des histones. Des complexes protéiques incluant des protéines reconnaissant les histones recrutent les méthyltransférases de l'ADN (Dnmts) pour méthylater l'ADN, induisant la liaison de complexes MBPs. (B) Modèle de modifications des histones dirigées par la méthylation de l'ADN. Les protéines se liant à l'ADN méthylé (MBPs) reconnaissent les CpG méthylés et recrutent des complexes corépresseurs comprenant des histones déacétylases (HDACs) ou des histones méthyltransférases (HMTs) pour former l'hétérochromatine, caractérisée par une hypoacétylation des histones et la méthylation d'H3K9. Ces deux mécanismes sont étroitement liés et contribuent à l'établissement et au maintien d'une structure chromatiniennne transcriptionnellement inactive.*

Plusieurs études suggèrent l'implication de la protéine UHRF1 dans les mécanismes de connexion entre la méthylation de l'ADN et les modifications des histones (Bronner et al., 2010; Hashimoto et al., 2009; Hashimoto et al., 2010). En effet, grâce à son domaine SRA, la protéine UHRF1 est capable de se lier à l'ADN méthylé (Arita et al., 2008; Avvakumov et al., 2008; Bostick et al., 2007; Hashimoto et al., 2008; Sharif et al., 2007). De plus, ses domaines PHD et TTD se lient spécifiquement à la queue N-terminale de l'histone H3 (Citterio et al., 2004) et seraient capables de reconnaître les modifications de l'histone au niveau d'H3K4 (Nady et al., 2011), d'H3K9 (Arita et al., 2012; Karagianni et al., 2008; Nady et al., 2011; Rottach et al., 2010; Walker et al., 2008; Xie et al., 2012) et d'H3R2 (Arita et al., 2012; Hu et al., 2011; Lallous et al., 2011; Rajakumara et al., 2011; Xie et al., 2012). Via son domaine RING qui possède une activité E3 ligase, UHRF1 serait impliquée dans l'ubiquitinylation des histones (Citterio et al., 2004; Jenkins et al., 2005; Karagianni et al., 2008). D'autre part, UHRF1 est connue pour s'associer avec une grande variété de facteurs épigénétiques tels que la Dnmt1 (Bostick et al., 2007; Sharif et al., 2007), l'histone méthyltransférase G9a (Kim et al., 2009b), et l'histone acétyltransférase Tip60 (Achour et al., 2009). Ainsi, grâce à ses différents domaines fonctionnels impliqués dans la lecture du code histone et la méthylation de l'ADN, UHRF1 pourrait constituer une plateforme de régulation des mécanismes épigénétiques.

L'ensemble de ces observations, témoigne de l'interconnexion existant entre les différents mécanismes de régulation épigénétique. D'autre part, elles révèlent l'existence de protéines qui peuvent simultanément contrôler le statut de méthylation des histones et de l'ADN afin de distinguer les différents états de la chromatine et permettre la régulation de son organisation en conséquence.

### 3.6. Dérégulation des modifications des histones

L'altération du système de modification des histones est liée à l'apparition de diverses maladies neurologiques. Ainsi, le syndrome de Rubinstein-Taybi est une maladie autosomique dominante caractérisée par un dysfonctionnement des HATs. Elle est causée selon le cas par une mutation des gènes CREBBP ("cAMP-response element binding protein") ou EP300 qui entraîne un retard mental associé à un aspect caractéristique de la face (Portela and Esteller, 2010; Urdinguio et al., 2009). Le syndrome de Coffin-Lowry est une maladie neuro-développementale liée au chromosome X associée à une mutation du gène RSK2, une sérine/thréonine kinase (Portela and Esteller, 2010). RSK2 permet l'activation de la transcription en remodelant la chromatine par phosphorylation directe d'H3S10 ou via son interaction avec l'HAT CREBBP. Cette maladie est caractérisée par un retard mental qui touche surtout les garçons, un retard de croissance, une hypotonie générale, et des déformations osseuses (Urdinguio et al., 2009). D'autre part, une hypoacétylation des histones est fréquemment associée à diverses maladies neurologiques telles que la sclérose latérale amyotrophique (SLA), la maladie de Parkinson, la maladie d'Huntington et l'ataxie de Friedreich (Urdinguio et al., 2009). D'autres altérations des profils de modifications des histones sont typiquement associées à la maladie d'Alzheimer et à l'épilepsie (Portela and Esteller, 2010).

Un rôle des profils de modification des histones a également été observé dans des maladies auto-immunes telles que la polyarthrite rhumatoïde et le diabète de type 1 dont les malades présentent des profils de modifications particuliers au niveau de gènes impliqués dans l'auto-immunité et l'inflammation (Portela and Esteller, 2010).

De même que pour la méthylation de l'ADN, la dérégulation des modifications d'histones est fréquemment corrélée à l'apparition de cancers. Ces dérégulations spécifiques seront discutées plus loin (voir p. 50).

#### 4. Les variants d'histones

En plus des quatre histones canoniques (H2A, H2B, H3 et H4), il existe plusieurs variants d'histones qui peuvent remplacer les histones conventionnelles au sein des nucléosomes. Ces variants d'histones diffèrent au niveau de leur queues (MacroH2A), dans la structure de leur domaine *histone-fold* (H2ABdb) ou simplement de quelques résidus (H3.3) (Li et al., 2007). Contrairement aux histones canoniques, ces variants sont exprimés en dehors de la phase S du cycle cellulaire et sont incorporés à la chromatine indépendamment de la réplication (Li et al., 2007). L'incorporation des variants d'histones au sein de la chromatine induit une restructuration des nucléosomes qui modifie leur positionnement et influence l'expression des gènes (Portela and Esteller, 2010; Santenard and Torres-Padilla, 2009 ; Sharma et al., 2010). En modifiant la structure de la chromatine, les variants d'histones jouent donc un rôle important dans l'établissement des propriétés locales de la chromatine dans laquelle ils sont enfouis (Santenard and Torres-Padilla, 2009). Cependant, les variants d'histones semblent conserver la plupart des sites de modifications post-traductionnelles des histones canoniques, permettant la reconnaissance des nucléosomes par les protéines de régulation de la chromatine (Li et al., 2007; McKittrick et al., 2004).

Les variants d'histones jouent des rôles majeurs dans divers mécanismes biologiques. Ainsi, le variant de l'histone H3, CenH3 est spécifique des centromères et essentiel pour l'assemblage du kinétochore (Talbert and Henikoff, 2010). En plus de CenH3, la plupart des eucaryotes possèdent un autre variant de l'histone H3, le variant H3.3, qui diffère de l'histone canonique de seulement quatre acides aminés. Il est retrouvé au niveau des gènes activement transcrits, au niveau des promoteurs, et associé à une activation de la transcription (Kamakaka and Biggins, 2005; Talbert and Henikoff, 2010). L'activation de la transcription peut déclencher son incorporation ou son éviction de la chromatine (Li et al., 2007; Talbert and Henikoff, 2010). Ce remplacement de l'histone canonique H3 par le variant H3.3 constituerait une marque de gènes actifs et favoriserait l'initiation de la transcription ultérieure ainsi que

l'établissement de nouvelles modifications post-traductionnelles de l'histone (Kamakaka and Biggins, 2005).

Parmi les variants de l'histone H2, H2A.Z est particulièrement bien étudié. Dans les cellules souches embryonnaires, H2A.Z colocalise avec les domaines bivalents de la chromatine et participerait à la régulation des gènes impliqués dans le développement (Creyghton et al., 2008; Sharma et al., 2010). Cette histone est enrichie au niveau de l'hétérochromatine et coopère avec HP1 $\alpha$  pour établir et maintenir la compaction de la chromatine en renforçant les interactions internucléosomiques (Fan et al., 2002) (Li and Reinberg, 2011; Talbert and Henikoff, 2010; Woodcock and Ghosh, 2010). D'autre part, le variant d'histone H2A.Z est enrichi au niveau de promoteurs de gènes où il contribuerait à l'activation de la transcription en protégeant l'ADN de la méthylation (Zilberman et al., 2008) (Portela and Esteller, 2010; Sharma et al., 2010). Ainsi, selon sa position dans la chromatine et ses modifications post-transcriptionnelles, le variant H2A.Z peut être associé à l'activation ou à la répression transcriptionnelle (Kamakaka and Biggins, 2005; Li et al., 2007; Sharma et al., 2010; Svtelis et al., 2009; Talbert and Henikoff, 2010; Zlatanova and Thakar, 2008). L'histone H2A.Z joue donc un rôle important dans la régulation de l'expression des gènes et la formation de l'hétérochromatine.

H2A.X, un autre variant de l'histone H2A, possède un motif C-terminal différent de son histone canonique qui peut être phosphorylée en réponse aux dommages de l'ADN (Li et al., 2007). Cette modification d'H2A.X permettrait le recrutement de protéines de réparation de l'ADN, d'enzymes de modifications des histones et de complexes de remodelage de la chromatine (Talbert and Henikoff, 2010). D'autre part, H2A.X serait requis pour la condensation et l'inactivation des chromosomes sexuels mâles lors de la méiose (Fernandez-Capetillo et al., 2003) et plus généralement dans la répression méiotique de la chromatine non-appariée (Talbert and Henikoff, 2010; Turner et al., 2005). H2A.X joue donc un rôle important dans le remodelage de la chromatine.

Le variant MacroH2A est enrichi au niveau du chromosome Xi (Costanzi and Pehrson, 2001). Son domaine *histone-fold* et sa longue queue C-terminale joueraient un rôle dans la répression transcriptionnelle des gènes par son action enzymatique (Ladurner, 2003) ou en bloquant stériquement la liaison des facteurs de transcription et des coactivateurs (Abbott et al., 2005) (Kamakaka and Biggins, 2005; Talbert and Henikoff, 2010).



Contrairement aux autres histones H2A, le variant H2A-Bdb ne possède pas de queue C-terminale. Cette propriété entraînerait une déstabilisation des nucléosomes qui faciliterait la transcription. En accord avec son rôle dans l'activation de la transcription, le variant H2A-Bdb est localisé au niveau du chromosome X activé et au niveau des chromosomes autosomes (Kamakaka and Biggins, 2005; Li et al., 2007).

Ainsi, l'inclusion de variants d'histones au sein des nucléosomes fournit un mécanisme épigénétique supplémentaire utilisé par les cellules pour modifier la structure de leur chromatine selon les besoins des divers processus cellulaires (Sharma et al., 2010).

## 5. Positionnement des nucléosomes

Les modifications des histones et/ou leur remplacement par des variants peut réguler la structure de la chromatine et l'expression des gènes en entraînant le repositionnement des nucléosomes au sein de la chromatine. En effet, les nucléosomes empêchent la liaison des facteurs de transcription et des coactivateurs à l'ADN et affectent tous les stades de la transcription en constituant une barrière physique à la transcription et à l'élongation (Portela and Esteller, 2010; Sharma et al., 2010). Le positionnement des nucléosomes est un mécanisme qui participe donc activement à la régulation de l'expression des gènes (Jiang and Pugh, 2009). En particulier, la position précise des nucléosomes au niveau des promoteurs joue un rôle important dans l'initiation de la transcription (Portela and Esteller, 2010). D'autre part, les régions dépourvues de nucléosomes aux extrémités 5' et 3' des gènes constitueraient des plateformes d'assemblage et de désassemblage nécessaires à la machinerie transcriptionnelle (Portela and Esteller, 2010; Sharma et al., 2010; Yuan et al., 2005). Ainsi, la perte d'un nucléosome immédiatement en amont du site d'initiation de la transcription est fortement corrélée à l'activation des gènes. Inversement, l'obstruction des sites d'initiation de la transcription par un nucléosome est associée à la répression transcriptionnelle (Cairns, 2009; Portela and Esteller, 2010; Schones et al., 2008; Sharma et al., 2010). La régulation de la transcription des gènes implique donc un équilibre dynamique entre l'empaquetage des séquences régulatrices au sein de la chromatine, et l'autorisation d'accès à des régulateurs de la transcription au niveau de ces séquences. Cet équilibre est régulé par des complexes de remodelage de la chromatine ATP-dépendant qui modifient l'accessibilité aux sites de régulation en éjectant les nucléosomes ou en les faisant glisser le long de l'ADN (Sharma et

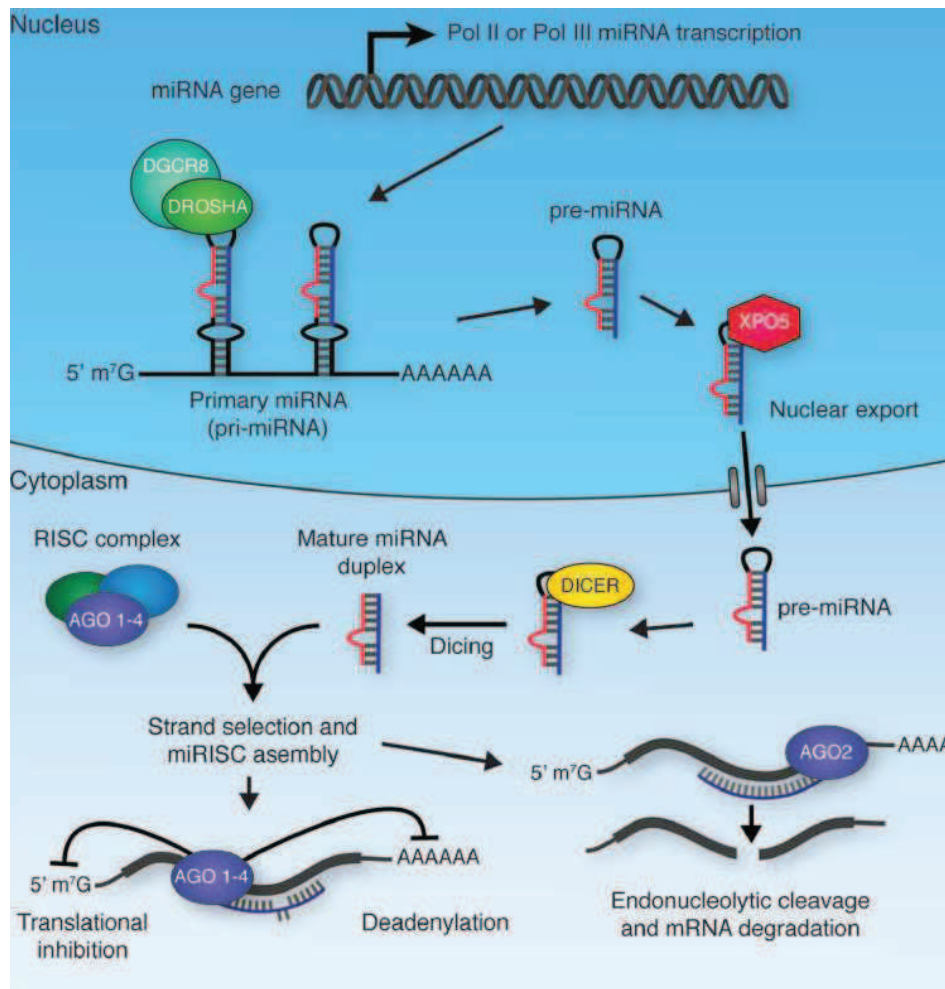
al., 2010; Smith and Peterson, 2005). Les modifications post-traductionnelles des histones participent activement à ce mécanisme en entraînant le recrutement des enzymes de remodelage de la chromatine. D'autre part, le remplacement des histones canoniques par des variants tels que H2A.Z, MacroH2A et H2A-Bbd joue également un rôle dans la régulation du positionnement des nucléosomes au sein de la chromatine. De plus, le positionnement des nucléosomes influence les profils de méthylation de l'ADN, les Dnmts ciblant préférentiellement l'ADN lié aux nucléosomes (Chodavarapu et al., 2010; Portela and Esteller, 2010). Ainsi, les acteurs du remodelage des nucléosomes communiquent avec la méthylation de l'ADN, les modifications des histones et les variants d'histones pour établir l'architecture de la chromatine et réguler l'expression des gènes (Sharma et al., 2010).

## 6. Les ARNs non-codant

Alors que seulement 1.2 % du génome humain code pour des protéines, près de 98 % des transcrits produits constituent des ARNs non-codants (ARNnc). Les ARNnc comprennent notamment les microARN (miARNs), les longs ARNs non-codant (lncARNs), les ARNs interagissant avec Piwi (ARNpi) et les larges ARNs non-codant intergéniques (LincARNs). Actuellement, ces ARNs s'imposent comme une nouvelle classe de gènes de régulation impliqués dans divers processus cellulaires tels que le développement, la différenciation cellulaire et le métabolisme (Mattick and Makunin, 2005; Wang and Chang, 2011; Zhang et al., 2012). D'autre part, ils semblent jouer un rôle important dans la régulation dynamique des chromosomes, la modification de la chromatine et la mémoire épigénétique, incluant l'empreinte génomique parentale, la méthylation de l'ADN et la répression transcriptionnelle.

Les miARNs sont de petits ARNs non-codant de 22 nt de long qui régulent l'expression des gènes par répression post-transcriptionnelle de gènes cibles. Ainsi, l'appariement de bases séquence-spécifiques d'un miARN avec la région non traduite 3'UTR ("untranslated region") d'un ARN messenger (ARNm), au sein du complexe RISC ("RNA-induced silencing complex"), conduit à la dégradation ou l'inhibition de la traduction de l'ARNm cible (Figure 14) (Ambros, 2004; Kim, 2005; Yang and Lai, 2011). Plus de 1000 miARNs existent dans le génome humain, modulant l'expression de 30 à 60 % des gènes

codant pour les protéines (Malumbres, 2012). Les miARNs sont exprimés de manière tissu-spécifique et contrôlent des processus biologiques variés tels que le développement, la prolifération cellulaire, la différenciation cellulaire, l'apoptose et la réponse au stress. D'autre part, en régulant des gènes suppresseurs de tumeur et des oncogènes, ils sont impliqués dans le développement de cancers, et leur expression peut se retrouver altérée dans les cellules tumorales (Iorio and Croce, 2012; Lee and Dutta, 2009).



*Figure 14 : Biogenèse des miARNs (Tollervey and Lunyak, 2012). Les transcrits primaires (pri-miARNs) sont générés par transcription sous l'action des ARNs polymérase II ou III. Le pri-miARN subit ensuite un clivage via l'action RNase III du complexe DROSHA/DGCR8 qui génère un pré-miARN d'environ 70 nt de long. XPO5 exporte ce pré-miARN dans le cytoplasme où il est davantage clivé par DICER afin de générer un duplex d'ARN double-brin mature. L'un des brins de ce duplex interagit avec l'une des quatre protéines Argonaute (AGO 1-4) pour former le complexe RISC qui peut moduler l'expression des gènes par inhibition de la traduction ou déadénylation d'un ARNm cible. Si le miARN s'apparie parfaitement à la séquence cible, un clivage endonucléolytique de l'ARNm a lieu sous l'action "slicer" d'AGO2.*

Les miARNs participent activement aux mécanismes épigénétiques en contrôlant l'expression de multiples gènes impliqués dans la régulation épigénétique tels que les Dnmts (Braconi et al., 2010; Duursma et al., 2008; Fabbri et al., 2007; Garzon et al., 2009), les histones déacétylases (Noonan et al., 2009) et les constituants des complexes répresseurs Polycomb (PCR1 et PCR2) (Cao et al., 2011; Iliopoulos et al., 2010; Wellner et al., 2009) (Malumbres, 2012). D'autre part, les miARNs ont tendance à cibler les régions promotrices de gènes non-méthylés, suggérant que ces deux mécanismes de régulation agissent de manière complémentaire afin d'inhiber l'expression des gènes (Su et al., 2011).

De récentes études suggèrent que, outre leur action de répression ou de dégradation des ARNm via le complexe RISC, les miARNs pourraient directement reconnaître des sites au niveau de promoteurs de gènes cibles et moduler l'expression des gènes en recrutant des enzymes de remodelage de la chromatine grâce à leur partenaires Argonaute (Malumbres, 2012). Ainsi, les miARNs peuvent induire ou réprimer l'expression de gènes spécifiques en entraînant des modifications épigénétiques sur leur promoteurs cibles (Huang et al., 2012; Kim et al., 2008; Place et al., 2008; Younger and Corey, 2011; Zardo et al., 2012) (Zhang et al., 2012).

Les miARNs constituent donc un mécanisme supplémentaire de régulation épigénétique qui agit en collaboration avec les enzymes de remodelage de la chromatine et de l'ADN pour réguler l'expression des gènes.

Plus de 1000 lncARNs, définis comme des molécules d'ARN de plus de 200 nt de long, ont été identifiés chez les mammifères (Guttman et al., 2009). La plupart des lncARNs sont transcrits par l'ARN polymérase II. A l'inverse des miARNs qui sont hautement conservés et impliqués dans la répression des gènes via un appariement de bases séquence-spécifiques ciblées, les lncARNs sont peu conservés et régulent l'expression des gènes par divers mécanismes encore mal définis (Wang and Chang, 2011). Les lncARNs sont impliqués dans la régulation post-transcriptionnelle des gènes en contrôlant des processus tels que la synthèse protéique, la maturation des ARNs et le transport (Bernstein and Allis, 2005; Whitehead et al., 2009). D'autre part, il est maintenant évident qu'ils jouent un rôle dans la répression transcriptionnelle en régulant l'architecture de la chromatine.

Les lncARNs présentent une expression tissu-spécifique et sont impliqués dans le développement embryonnaire, la différenciation cellulaire (Dinger et al., 2008), l'inactivation

du chromosome X (Pontier and Gribnau, 2011) et l'établissement de l'empreinte génomique parentale (Mohammad et al., 2009) (Wang and Chang, 2011; Zhang et al., 2012). D'autre part, en réagissant aux stimuli de l'environnement, ils constitueraient un élément clé de la régulation de la transcription dans des processus tels que la réponse aux dommages de l'ADN et la régulation du cycle cellulaire (Hung et al., 2011).

Les lncARNs possèdent des structures variées et peuvent réguler l'expression des gènes selon des mécanismes épigénétiques distincts (Wang and Chang, 2011; Zhang et al., 2012). Par exemple, les lncARNs peuvent constituer des leurres et lier des protéines cibles telles que des facteurs de transcription, des protéines de remodelage ou d'autres facteurs de régulation, les empêchant d'agir sur leur cibles originelles et donc d'exécuter leur fonction (Wang and Chang, 2011). D'autre part, les lncARNs peuvent se lier à des protéines et leur servir de guide en dirigeant leur localisation au niveau de cibles spécifiques. Par ce mécanisme, les lncARNs peuvent donc induire des variations de l'expression de gènes proches dans la séquence (action en *cis*) ou éloignés, sur un autre chromosome (action en *trans*) (Wang and Chang, 2011). Enfin, les lncARNs serviraient d'adaptateur et de plateforme de centralisation permettant la liaison simultanée de multiples partenaires effecteurs au niveau d'un même site génomique (Wang and Chang, 2011).

Ainsi, les lncARNs sont impliqués dans des mécanismes biologiques essentiels et constituent un processus épigénétique clé capable d'intégrer les informations de l'environnement et de moduler l'expression des gènes de multiples manières selon leur structure et leur partenaire.

## B. Épigénétique et cancer

Traditionnellement, c'est l'accumulation des mutations génétiques qui est considérée comme la cause principale des maladies cancéreuses. Cependant, la découverte de profondes altérations des mécanismes épigénétiques accompagnant ces mutations a révélé leur rôle déterminant dans l'initiation et la progression de la maladie (Rodriguez-Paredes and Esteller, 2011; Sandoval and Esteller, 2012; You and Jones, 2012). Ainsi, l'épigénome des cellules cancéreuses est caractérisé par un changement global de la méthylation de l'ADN, l'altération des profils de modifications des histones, et la dérégulation de l'expression des enzymes de

remodelage de la chromatine. Ces caractéristiques étant typiquement observées dans les néoplasies bénignes et les premiers stades de tumeurs, il est clair que les altérations épigénétiques précèdent les événements classiques de transformation tels que les mutations des gènes suppresseurs de tumeur et/ou des proto-oncogènes et l'instabilité génomique, qui entraînent la perte ou le gain de fonctions et des profils d'expression anormaux dans les cellules cancéreuses. L'altération des mécanismes de régulation épigénétiques survient par mutation, délétion ou dérégulation de l'expression de leurs protagonistes, entraînant des profils d'expression aberrants et une instabilité génomique prédisposant au cancer. D'autre part, ces "épimutations" peuvent promouvoir l'initiation du cancer en réprimant l'allèle encore actif d'un gène suppresseur de tumeur déjà muté. Ainsi, l'épigénétique et la génétique sont des mécanismes qui coopèrent dans l'initiation et la progression des maladies cancéreuses.

## 1. Hyperméthylation locale versus hypométhylation globale

L'altération des profils de méthylation de l'ADN est l'une des dérégulations épigénétiques les plus caractéristiques des cellules cancéreuses (Rodriguez-Paredes and Esteller, 2011; Sandoval and Esteller, 2012). Ainsi, au cours de la tumorigénèse, on observe progressivement une hypométhylation globale du génome avec une perte de 20 à 60 % des cytosines méthylées, accompagnée d'une hyperméthylation ciblée des îlots CpG au niveau de certains promoteurs (Fernandez et al., 2012a). Les données épigénomiques accumulées depuis plusieurs années à partir de divers classes de néoplasies humaines révèlent l'existence d'un profil d'hyperméthylation unique définissant chaque type de cancer (Ballestar and Esteller, 2008; Costello et al., 2000; Esteller et al., 2001; Fernandez et al., 2012a).

Dans les cellules cancéreuses, l'hypométhylation de l'ADN est principalement observée au niveau de régions génomiques pauvres en gènes telles que les séquences répétées, favorisant l'instabilité génomique, les translocations, la disruption génique et la réactivation de séquences endoparasites (Portela and Esteller, 2010; Rodriguez-Paredes and Esteller, 2011; Sandoval and Esteller, 2012). Cependant, l'hypométhylation de l'ADN peut occasionnellement avoir lieu au niveau du site d'initiation de la transcription de promoteurs ne contenant pas d'îlots CpG. L'hypométhylation de ces sites pourrait être impliquée dans l'activation de proto-oncogènes. D'autre part, la méthylation spécifique de ces sites jouant un rôle fondamentale dans la différenciation cellulaire et l'établissement de l'empreinte génomique parentale, leur



hypométhylation entraînerait la perte de l'identité cellulaire des cellules transformées et/ou la suppression de l'empreinte, favorisant ainsi l'apparition de cancers (Fernandez et al., 2012a; Lim and Maher, 2010; Rodriguez-Paredes and Esteller, 2011; Sandoval and Esteller, 2012). La dérégulation fréquente des profils de méthylation des "CpG island shores" observées dans les cellules cancéreuses renforce davantage l'hypothèse d'altérations épigénétiques affectant la différenciation tissu-spécifique et augmentant le risque de cancer (Irizarry et al., 2009).

L'hyperméthylation de l'ADN observée dans le cancer est plus localisée et principalement présente au niveau des îlots CpG (78 %) (Fernandez et al., 2012a). Ainsi, les cellules cancéreuses sont caractérisées par une hyperméthylation aberrante des promoteurs des gènes suppresseurs de tumeurs généralement associée à la répression des gènes (Fernandez et al., 2012a) (Esteller, 2008). Selon le type de tumeur on observe l'inactivation de gènes impliqués dans la régulation du cycle cellulaire (RB1, p16<sup>INK4A</sup>), l'apoptose (DAPK1, RASSF1A), la réparation de l'ADN (MLH1, BRCA1) ou d'autres processus cellulaires (Esteller, 2008; Rodriguez-Paredes and Esteller, 2011). Les études épigénomiques révèlent ainsi l'existence de 500 promoteurs de gènes hyperméthylés de manière aberrante dans les tumeurs (Fernandez et al., 2012a).

L'apparition de profils de méthylation aberrants dans le cancer pourrait être expliquée par les altérations des nombreux acteurs de la méthylation de l'ADN. Ainsi, on observe la surexpression des Dnmts et leur mutation dans plusieurs types de cancers (Dawson and Kouzarides, 2012; Rodriguez-Paredes and Esteller, 2011; Sandoval and Esteller, 2012; You and Jones, 2012). D'autre part, une étude récente révèle une corrélation entre l'hypométhylation globale du génome et la diminution significative des complexes Dnmt1/PCNA/UHRF1 due à un défaut d'interactions entre les protéines (Hervouet et al., 2010a). Des mutations de la protéine TET2, responsable de la conversion de la 5mC en 5hmC, on également été rapportées (Dawson and Kouzarides, 2012; Rodriguez-Paredes and Esteller, 2011; You and Jones, 2012). L'altération des miARNs observés dans certains types de cancers pourrait également participer à la dérégulation des acteurs de la méthylation (Sandoval and Esteller, 2012).



## 2. Modifications des histones dans le cancer

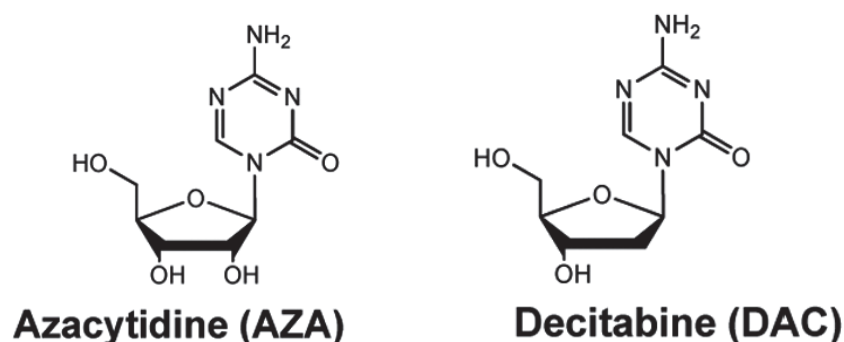
Au cours de la tumorigenèse, outre l'altération des profils de méthylation de l'ADN, on observe des changements dans les modifications post-traductionnelles des histones qui affecteraient le recrutement des facteurs de transcription et des constituants de la machinerie transcriptionnelle, contribuant à l'expression aberrante de gènes (Dawson and Kouzarides, 2012; Rodriguez-Paredes and Esteller, 2011; Sandoval and Esteller, 2012; You and Jones, 2012). Ainsi, dans les tumeurs primaires, on observe fréquemment une réduction globale d'H4K20me3 et d'H4K6Ac au niveau des séquences répétées hypométhylées (Fraga et al., 2005) (Portela and Esteller, 2010; Rodriguez-Paredes and Esteller, 2011; Sandoval and Esteller, 2012). De nombreuses études révèlent que l'altération des profils de modification des histones observée dans les cellules cancéreuses est principalement due à une expression aberrante des acteurs de l'acétylation et de la méthylation des histones (Dawson and Kouzarides, 2012; Portela and Esteller, 2010; Rodriguez-Paredes and Esteller, 2011; Sandoval and Esteller, 2012; You and Jones, 2012). Chaque tumeur peut donc être distinguée par un profil d'expression spécifique des enzymes de modification des histones (Ozdogan et al., 2006). D'autre part, selon le type de cancer, des altérations sont observées dans les protéines qui éliminent (HDM, HDAC) ou reconnaissent (la protéine à chromodomaine HP1 qui reconnaît les lysines méthylées, la protéine Brd4 qui reconnaît les lysines acétylées) les modifications spécifiques des histones (Dawson and Kouzarides, 2012; Rodriguez-Paredes and Esteller, 2011; Sandoval and Esteller, 2012; You and Jones, 2012). D'autres mutations fréquemment observées dans les cancers concernent les protéines constituant les complexes répresseurs Polycomb (BMI1, EZH2) qui contrôlent l'accessibilité de la machinerie transcriptionnelle au niveau d'éléments de régulation de gènes et qui jouent un rôle majeur dans le développement (Dawson and Kouzarides, 2012; Sandoval and Esteller, 2012; You and Jones, 2012).

## 3. Thérapie épigénétique

A l'inverse des mutations génétiques, les "épimutations" sont réversibles. Les acteurs de la régulation épigénétique font donc l'objet d'intenses recherches afin de découvrir de nouveaux agents thérapeutiques capables de rétablir des profils épigénétiques normaux dans

les cellules cancéreuses (Arrowsmith et al., 2012; Dawson and Kouzarides, 2012; Rodriguez-Paredes and Esteller, 2011; You and Jones, 2012).

Actuellement, quatre molécules de régulation épigénétique ont été approuvées par la FDA ("Food and Drug Administration") dans le traitement des cancers. Ainsi, deux inhibiteurs des Dnmts, l'azacytidine (5-azacytidine, Vidaza<sup>®</sup>) et la décitabine (5-aza-2'-déoxycytidine, Dacogen<sup>®</sup>) sont déjà sur le marché (Figure 15) (Dawson and Kouzarides, 2012; Rodriguez-Paredes and Esteller, 2011). Ces analogues nucléosidiques sont incorporés à l'ADN à la place du résidu cytosine lors de la réplication. La présence d'un groupement nitrogène à la position 5 de ces bases modifiées induit la formation d'un complexe covalent irréversible entre la base et la Dnmt1, entraînant la dégradation de la Dnmt1 par le protéasome (Ghoshal et al., 2005). D'autre part, l'absence de cytosine rend impossible la méthylation de l'ADN par les diverses Dnmts. Il en résulte une diminution significative de la méthylation globale de l'ADN. Bien que ces agents thérapeutiques aient montré des résultats mitigés sur les tumeurs solides, ils ont prouvé leur efficacité dans les syndromes myélodysplasiques. Leur manque de spécificité entraîne une toxicité mais l'administration de doses faibles de ces agents thérapeutiques suffit à établir un bénéfice thérapeutique durable (Tsai et al., 2012). L'inhibition des Dnmts apparaît comme une stratégie thérapeutique prometteuse dans le traitement des cancers et la recherche de nouveaux inhibiteurs plus spécifiques et moins toxiques est toujours à l'étude (Medina-Franco and Caulfield, 2011).



*Figure 15 : Structures moléculaires de l'azacytidine et de la décitabine.*

Les deux autres molécules de régulation épigénétique existant sur le marché comme agents thérapeutiques anticancéreux sont des inhibiteurs des histones déacétylases (HDACs), le vorinostat (acide suberoylanilide hydroxamique, Zolinza<sup>®</sup>) et la romidepsin (FK-228,

Istodax<sup>®</sup>) (Dawson and Kouzarides, 2012; Rodriguez-Paredes and Esteller, 2011). Ces agents thérapeutiques sont caractérisés par la présence d'un domaine liant les métaux qui bloque la molécule de zinc trouvée dans le site actif des HDACs de classes I, II et IV. Ils sont capables d'annuler certaines des répressions géniques aberrantes observées dans les cancers et induisent l'arrêt du cycle cellulaire, la différenciation cellulaire, et l'apoptose dans les cellules cancéreuses (Federico and Bagella, 2011). Ainsi, ces deux molécules ont récemment été approuvées dans le traitement de patients atteints de lymphomes T cutanés (Olsen et al., 2007; Piekarczyk et al., 2009). Les résultats encourageants obtenus sur les inhibiteurs des HDACs ont fait d'eux de bons candidats pour le développement de nouvelles thérapies anticancéreuses et plusieurs nouveaux inhibiteurs sont actuellement sous investigation pour une utilisation clinique dans plusieurs types de cancers (Federico and Bagella, 2011).

En plus des inhibiteurs des Dnmts et des HDACs, d'autres acteurs de la régulation épigénétique font l'objet d'intenses recherches. On observe ainsi des résultats préliminaires encourageants pour des inhibiteurs des HATs, des HMTs, des HDMs et des kinases, ainsi que des protéines impliquées dans la lecture du code histone telles que les protéines à bromodomaine, qui reconnaissent les lysines acétylées, et les protéines reconnaissant les lysines méthylées (Arrowsmith et al., 2012; Dawson and Kouzarides, 2012; Rodriguez-Paredes and Esteller, 2011; You and Jones, 2012).

La présence simultanée de multiples aberrations génétiques et épigénétiques dans les cancers suggère que les médicaments épigénétiques combinés aux chimiothérapies standards joueraient un rôle complémentaire bénéfique dans le traitement des cancers (Juergens et al., 2011; Matei and Nephew, 2010). Ainsi, ces agents thérapeutiques deviendront certainement dans le futur un arsenal thérapeutique essentiel pour combattre le cancer.

### C. La protéine humaine hUHRF1 et la régulation épigénétique

A l'origine, la protéine humaine hUHRF1 pour "Ubiquitin-like PHD Ring Finger 1", appelée alors ICBP90 pour "Inverted CCAAT box Binding Protein of 90 kDa", avait été découverte en 2000 par l'équipe de Christian Bronner grâce à la technique du simple hybride en tant que protéine se liant à la deuxième boîte inversée CCAAT (ICB2) du promoteur du gène de la *topoisomérase II $\alpha$*  (Hopfner et al., 2000). Le gène codant pour *hUHRF1* est

localisé dans la région chromosomique 19p13.3 et code pour une protéine nucléaire de 793 acides aminés de poids moléculaire théorique de 89,756 kDa (Hopfner et al., 2001). Cette protéine est impliquée dans divers processus cellulaires tels que la régulation du cycle cellulaire, la réponse aux dommages de l'ADN et la transmission du code épigénétique. D'autre part, elle se trouve surexprimée dans différents types de cancers tels que le cancer du sein, de la vessie, du foie, et du pancréas, ce qui fait d'elle une cible thérapeutique potentielle intéressante.

## 1. La famille UHRF

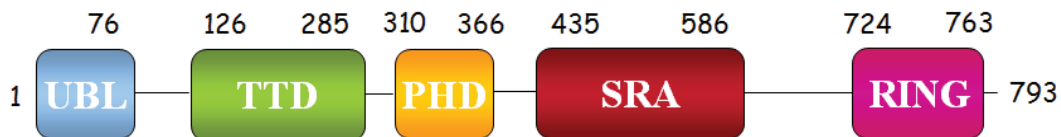
La famille UHRF1 comprend trois protéines humaines: hUHRF1 (ou ICPB90), hUHRF2 (ou NIRF pour "Np95/ICBP90 Ring Finger") et hUHRF3 (ou ICPB55); et trois protéines d'origine murine: mUHRF1 (ou Np95), mUHRF2 (ou Np97) et mUHRF3 (ou Np55). Les alignements de séquences d'acides aminés des différentes protéines de la famille UHRF révèlent 74 % d'homologie entre hUHRF1 et mUHRF1, et 53 % d'homologie entre hUHRF1 et hUHRF2 (Bronner et al., 2007a).

Malgré son homologie de structure avec hUHRF1, hUHRF2 semble posséder des fonctions distinctes (Bronner et al., 2007a; Mori et al., 2012). D'une part, lors de la différenciation, hUHRF1 et hUHRF2 présentent des profils d'expression opposés, suggérant des rôles fonctionnels différents dans ce processus cellulaire. Ainsi, contrairement à hUHRF1, hUHRF2 est exprimé de manière ubiquitaire dans presque tous les tissus, surexprimé lors de la différenciation cellulaire, et exprimé dans les tissus différenciés tels que le cerveau, les poumons et le foie (Pichler et al., 2011). D'autre part, hUHRF1 et hUHRF2 présentent des effets opposés dans le contrôle du cycle cellulaire, la surexpression de hUHRF2 entraînant l'arrêt du cycle cellulaire en phase G1 (Li et al., 2004). Enfin, hUHRF2 est incapable de recruter la Dnmt1 au niveau de la fourche de réplication lors de la phase S du cycle cellulaire et ne permet pas de restaurer le défaut de méthylation de l'ADN dans les cellules embryonnaires défectueuses en hUHRF1 (Zhang et al., 2011). Ainsi, à l'inverse de hUHRF1, hUHRF2 semble se comporter comme un gène suppresseur de tumeur (Mori et al., 2011). En effet, le gène hUHRF2 est fréquemment perdu dans plusieurs types de tumeurs et une microdélétion ciblant UHRF2 est observée de manière récurrente dans les carcinomes du poumon (Mori et al., 2011). Ces observations suggèrent que hUHRF2 occupe une position

importante dans la régulation du cycle cellulaire et confirment qu'il pourrait constituer un gène suppresseur de tumeur dont les aberrations contribueraient à la tumorigenèse (Mori et al., 2011). Contrairement aux autres protéines de la famille UHRF, les protéines hUHRF3 et mUHRF3 ne possèdent pas de domaine UBL. L'absence de ce domaine fonctionnel suggère des rôles distincts pour mUHRF3 et hUHRF3. Cependant, leur rôles, ainsi que celui de la protéine mUHRF2, n'ont pas encore été étudiés (Bronner et al., 2007a).

## 2. Architecture de la protéine hUHRF1

La protéine UHRF1 possède cinq domaines structuraux ayant des fonctions distinctes (Figure 16) (Bronner et al., 2007a): le domaine UBL (ou NIRF-N) pour "Ubiquitin-like domain", le domaine TTD pour "cryptic Tandem Tudor Domain", le domaine PHD pour "Plant Homeo Domain", le domaine SRA pour "Set and Ring Associated domain", et le domaine RING pour "Really Interesting New Gene domain".



*Figure 16 : Domaines structuraux de hUHRF1. UBL, "Ubiquitin-like domain"; TTD, "cryptic Tandem Tudor Domain"; PHD, "Plant Homeo Domain"; SRA, "Set and Ring Associated domain"; RING, "Really Interesting New Gene domain".*

### 2.1. Le domaine UBL ("Ubiquitin-like domain")

Le domaine UBL (ou NIRF-N) est situé à l'extrémité N-terminale de la protéine hUHRF1 et défini par les acides aminés 1 à 76 (Bronner et al., 2007a). Bien que la fonction de ce domaine reste à définir, son homologie de séquence de 35 % avec l'ubiquitine suggère qu'il pourrait mimer l'ubiquitine constitutive et interagir avec le protéasome. Ainsi, grâce à ce domaine, la protéine hUHRF1 pourrait jouer un rôle dans la régulation du cycle cellulaire, la dégradation de protéines et la transcription de gènes. L'ubiquitinylation des histones via le domaine RING de la protéine hUHRF1 et sa liaison potentielle avec le protéasome via son

domaine UBL constituerait un mécanisme de régulation des modifications post-traductionnelles des histones en permettant la dégradation des histones ubiquitinylées via le protéasome.

## 2.2. Le domaine TTD ("cryptic Tandem Tudor Domain")

Entre les acides aminés 126 et 285 de la protéine hUHRF1, un domaine TTD constitué de deux sous-domaines a été identifié en 2008 (Walker et al., 2008). Ce domaine TTD, en reconnaissant spécifiquement la queue N-terminale de l'histone H3, jouerait un rôle dans la lecture du code histone. Ainsi, le domaine TTD de UHRF1 se lie préférentiellement à H3K9me3 (Hu et al., 2011), la structure cristallographique du complexe TTD/histone H3 révélant la formation d'une poche hydrophobe grâce aux deux sous-domaines de TTD pouvant s'adapter à l'histone H3, la lysine 9 triméthylée étant spécifiquement reconnue au sein du premier sous-domaine via une cage aromatique (Rottach et al., 2010; Walker et al., 2008). Deux études récentes suggèrent que l'association du domaine TTD au domaine PHD potentialiserait la liaison d'UHRF1 à H3K9me3 (Arita et al., 2012; Xie et al., 2012). D'autre part, dans le complexe TTD/histone H3, le sillon formé entre les deux sous-domaines de TTD permettrait la lecture simultanée de la lysine 4 non-méthylée (H3K4me0) et d'H3K9me3 (Nady et al., 2011). L'existence de protéines pouvant reconnaître simultanément H3K9me3 et H3K4me0 est cohérente avec l'observation récente de méthyltransférases des histones (SETB1, Suv39h1, G9a et GLP) capables de méthyler préférentiellement la lysine 9 des histones H3 non-méthylées sur la lysine 4 (Binda et al., 2010). D'autre part, confirmant la liaison préférentielle du domaine TTD, UHRF1 est enrichie au niveau des nucléosomes contenant H3K4me0/K9me3 (Bartke et al., 2010). La marque H3K4me0/K9me3 étant caractéristique de l'hétérochromatine, le domaine TTD pourrait participer à la connexion entre la méthylation de l'ADN et les modifications des histones en permettant la méthylation de l'ADN au niveau de ces régions spécifiques de la chromatine, favorisant la répression des gènes en stabilisant l'hétérochromatine (Bartke et al., 2010; Bronner et al., 2010; Hashimoto et al., 2009; Hashimoto et al., 2010). Supportant cette hypothèse une étude récente montre ainsi que la liaison du domaine TTD à l'histone H3 méthylée sur la lysine 9 est nécessaire à la méthylation de maintenance de l'ADN (Rothbart et al., 2012).

### 2.3. Le domaine PHD ("Plant Homeo Domain")

Le domaine PHD de hUHRF1 est défini par les acides aminés 310 à 366. Cependant, le doigt de zinc précédant la région conventionnelle du domaine PHD semble jouer un rôle important dans les fonctions associées au domaine PHD (Lallous et al., 2011; Xie et al., 2012). Selon l'équipe de Jacobsen, le domaine PHD serait responsable de la liaison de UHRF1 avec la Dnmt1 (Bostick et al., 2007). D'autre part, le domaine PHD de UHRF1 serait impliqué dans la réorganisation des domaines péri-centromériques de la chromatine (Karagianni et al., 2008; Papait et al., 2008). Ces régions correspondent à de l'hétérochromatine associée de manière très compacte et organisée en chromocentres. Le recrutement de UHRF1 au niveau de l'hétérochromatine péri-centromérique serait possible grâce à son domaine PHD qui permettrait la réplication de ces régions en désagrégeant les chromocentres pour faciliter l'accès à la machinerie de réplication et aux protéines de modification de la chromatine au niveau de l'ADN. De manière similaire au domaine TTD, le domaine PHD serait impliqué dans la lecture du code histone en reconnaissant spécifiquement les modifications chimiques de l'histone H3 (Hu et al., 2011; Lallous et al., 2011; Rajakumara et al., 2011; Xie et al., 2012). Bien que le domaine PHD présente une préférence pour H3K4 non méthylée (H3K4me0) (Hu et al., 2011; Xie et al., 2012), sa structure cristallographique montre qu'il est capable de s'accommoder à H3K4me3 sans compromettre la formation du complexe PHD/histone H3 (Lallous et al., 2011; Xie et al., 2012). Ainsi, le domaine PHD semble reconnaître spécifiquement l'histone H3 non modifiée principalement via la reconnaissance de l'arginine 2 non méthylée (H3R2me0), la méthylation de ce résidu pouvant diminuer jusqu'à 20 fois la liaison du domaine PHD (Hu et al., 2011; Lallous et al., 2011; Rajakumara et al., 2011; Xie et al., 2012). Bien que la liaison spécifique de PHD à H3K9me3 ait été initialement identifiée (Karagianni et al., 2008), plusieurs études récentes démontrent que la liaison du domaine PHD à l'histone H3 a lieu indépendamment de la méthylation d'H3K9 (Hu et al., 2011; Lallous et al., 2011; Nady et al., 2011; Papait et al., 2008; Rajakumara et al., 2011; Xie et al., 2012). Ainsi, ce serait l'association des domaines PHD et TTD qui permettrait de reconnaître spécifiquement la marque H3K9me3 (Xie et al., 2012). D'autre part, l'association de ces deux domaines permettrait la reconnaissance simultanée d'H3K9me3 et d'H3R2me0 (Arita et al., 2012). Ce sont les résidus reliant les deux domaines qui permettent la formation d'une poche reconnaissant l'histone H3, leur phosphorylation modulant la position relative des modules de reconnaissance d'H3R2 et d'H3K9. Ainsi, en reconnaissant spécifiquement les



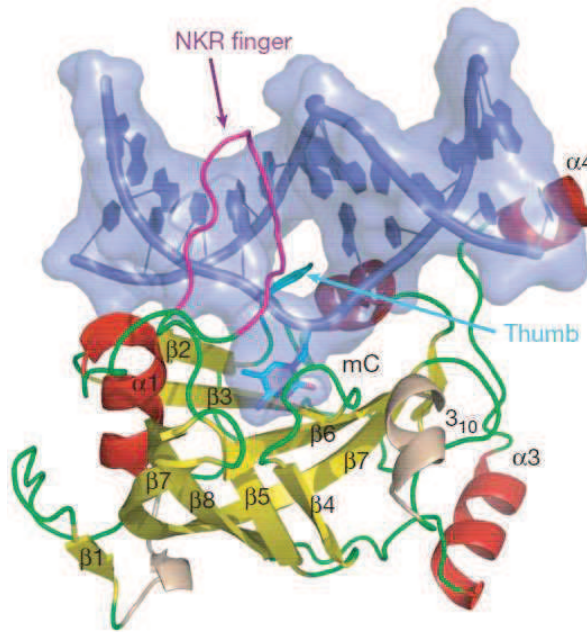
modifications de l'histone H3, les domaines TTD et PHD semblent jouer un rôle essentiel dans la connexion entre la méthylation de l'ADN et la modification des histones.

#### 2.4. Le domaine SRA ("Set and Ring Associated domain")

Le domaine SRA de hUHRF1 est défini par les acides aminés 435 à 586. Ce domaine est fréquemment observé parmi les protéines des plantes mais présent chez les vertébrés uniquement dans la famille UHRF.

La fonction majeure du domaine SRA est de reconnaître l'ADN hémi-méthylé, contribuant à la localisation nucléaire de la protéine UHRF1 (Arita et al., 2008; Avvakumov et al., 2008; Bostick et al., 2007; Frauer et al., 2011; Hashimoto et al., 2008; Qian et al., 2008; Rottach et al., 2010; Sharif et al., 2007; Unoki et al., 2004). Bien que deux études aient initialement révélé l'interaction du domaine SRA avec l'histone H3 (Citterio et al., 2004; Karagianni et al., 2008), plusieurs travaux récents réfutent cette première hypothèse (Hu et al., 2011; Nady et al., 2011; Rajakumara et al., 2011), suggérant que c'est l'interaction du domaine SRA avec l'ADN méthylé qui entraînerait sa présence au niveau des histones. En effet, on a vu que ce sont les domaines TTD et PHD qui sont impliqués dans la liaison de UHRF1 aux histones. Ces deux domaines, associés au domaine SRA, seraient responsables de la localisation nucléaire de la protéine UHRF1 (Rottach et al., 2010).

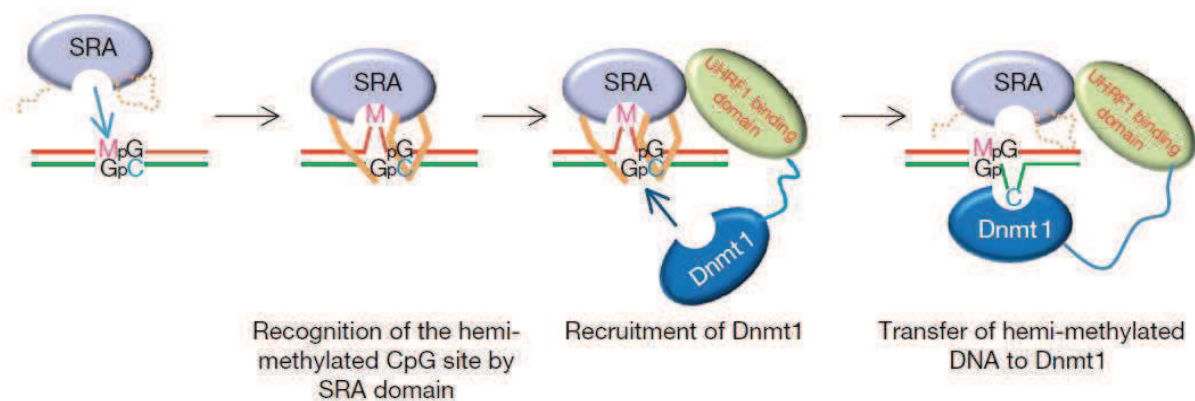
Grâce à son interaction simultanée avec l'ADN méthylé et de nombreux acteurs de la régulation épigénétique, le domaine SRA de UHRF1 jouerait un rôle central dans la connexion entre la méthylation de l'ADN et les modifications des histones. En effet, via son interaction à l'ADN, le domaine SRA permettrait de recruter de nombreuses protéines de régulation épigénétique au niveau de la chromatine telles que la Dnmt1 (Achour et al., 2008; Felle et al., 2011), HDAC1 (Unoki et al., 2004), et l'histone méthyltransférase G9a (Kim et al., 2009b). Ces protéines participent à la formation d'un gros complexe macromoléculaire permettant la régulation de la transcription et qui pourrait être responsable de la transmission du code épigénétique (Bronner et al., 2007b; Kim et al., 2009b).



*Figure 17 : Structure du domaine SRA de UHRF1 en complexe avec l'ADN hémi-méthylé (Avvakumov et al., 2008). Le domaine SRA est représenté avec ses boucles flexibles, ses feuillettes  $\beta$  et ses hélices  $\alpha$  en vert, jaune et rouge, respectivement. Le duplex d'ADN est montré dans une représentation de surface avec les atomes et les bases de l'hélice colorés en bleu. La cytosine méthylée basculée est annotée "mC". Le doigt NKR (magenta) et le pouce (cyan) de la protéine prennent la place de la cytosine basculée dans l'espace libéré dans l'ADN.*

De nombreuses études décrivent l'affinité préférentielle du domaine SRA pour l'ADN hémi-méthylé (ADN méthylé sur l'un des deux brins uniquement) (Arita et al., 2008; Avvakumov et al., 2008; Bostick et al., 2007; Frauer et al., 2011; Greiner et al., en préparation; Hashimoto et al., 2008; Qian et al., 2008; Rottach et al., 2010). Ainsi, via son domaine SRA, la protéine UHRF1 participerait au maintien des profils de méthylation de l'ADN en recrutant la Dnmt1 au niveau des sites CpG hémi-méthylés générés après la réplication. La résolution de la structure tridimensionnelle du domaine SRA (Delagoutte et al., 2008; Qian et al., 2008) et du complexe SRA/ADN (Arita et al., 2008; Avvakumov et al., 2008; Hashimoto et al., 2008) a permis de proposer un modèle moléculaire de reconnaissance de l'ADN hémi-méthylé. Le domaine SRA de la protéine UHRF1 se comporte comme une main dont le doigt NKR (Asp 489, Lys 490, Arg 491) et le pouce permettent le basculement mécanique de la cytosine méthylée dans le grand sillon de l'ADN (Figure 17) (Arita et al., 2008; Avvakumov et al., 2008; Hashimoto et al., 2008). Ce mécanisme de basculement de base ("base-flipping") est un processus largement utilisé par les enzymes de modification des nucléotides telles que les méthyltransférases de l'ADN (Cheng and Roberts, 2001;

Klimasauskas et al., 1994), les enzymes de réparation de l'ADN (Parker et al., 2007; Yang et al., 2008) et les enzymes de modification des ARNs (Lee et al., 2005). Cependant, le domaine SRA est le premier domaine protéique connu capable de se lier spécifiquement à l'ADN et de basculer une base sans activité enzymatique. Alors que les enzymes adoptent ce mécanisme afin d'avoir accès à la base d'ADN qu'elles vont modifier chimiquement, le domaine SRA, dont la séquence de reconnaissance est de seulement deux paires de bases, profite de ce mécanisme probablement pour augmenter sa surface d'interaction avec l'ADN et prévenir sa diffusion linéaire loin du site ciblé sur l'ADN (Hashimoto et al., 2009). Ainsi, ce mécanisme de reconnaissance des sites hémi-méthylés par le domaine SRA, permettrait à la protéine UHRF1 de s'ancrer à des endroits stratégiques sur l'ADN, facilitant le recrutement de la Dnmt1 au niveau des sites CpG hémi-méthylés. Dans le modèle proposé par l'équipe de Shirakawa (Arita et al., 2008), le complexe ternaire serait transitoire et la liaison de la Dnmt1 au domaine SRA déstabiliserait le complexe SRA/ADN, permettant le transfert du site CpG hémi-méthylé vers la Dnmt1 qui induirait le basculement de la cytosine non-méthylée sur le brin "fille" dans le site actif de l'enzyme, permettant ainsi la méthylation de l'ADN sur le brin nouvellement synthétisé (Figure 18). Le domaine SRA de UHRF1 et l'enzyme Dnmt1 travailleraient donc en collaboration afin de transmettre les profils de méthylation lors de la réplication de l'ADN.



*Figure 18 : Modèle schématique du mécanisme de maintenance des profils de méthylation de l'ADN après réplication (Arita et al., 2008). Le domaine SRA de UHRF1 et son partenaire Dnmt1 fonctionnent en coopération pour maintenir la méthylation des dinucléotides CpG. M représente la cytosine méthylée, le brin d'ADN nouvellement synthétisé après réplication de l'ADN est représenté en vert, le brin rouge représentant le brin parental.*

Les mécanismes par lesquels la protéine UHRF1 se lie préférentiellement à l'ADN hémi-méthylé restent à éclaircir. Une étude récente révèle que l'ajout d'un groupement

méthyle à la cytosine induit une redistribution des électrons autour du cycle de la cytosine réduisant son moment dipolaire (Bianchi and Zangi, 2012). Par conséquent, l'ADN non-méthylé est plus stable en solution que l'ADN hémi-méthylé. En outre, l'énergie d'interaction de la protéine UHRF1 avec l'ADN hémi-méthylé est plus forte qu'avec l'ADN non-méthylé. Ainsi, les variations de la structure électronique de la cytosine méthylée déstabilisent son état non lié et favorise l'interaction de l'ADN hémi-méthylé avec la protéine UHRF1. D'autre part, les études de la structure tridimensionnelle du complexe SRA/ADN révèlent que la poche de liaison de la protéine s'ajuste exactement à la cytosine méthylée (Arita et al., 2008; Avvakumov et al., 2008; Hashimoto et al., 2008). Les bases entourant la cytosine basculée seraient alors stabilisées par l'insertion du doigt NKR au sein de l'hélice d'ADN, favorisant l'ancrage du domaine SRA au niveau des sites CpG hémi-méthylés. La liaison du domaine SRA à l'ADN bi-méthylé serait quant à elle déstabilisée par la présence du groupement méthyle sur les deux cytosines du site CpG. L'encombrement stérique créé déséquilibrerait l'état conformationnel du doigt NKR au sein de l'ADN et serait responsable de la diminution d'affinité de UHRF1 observée pour l'ADN bi-méthylé.

Bien que la plupart des études rapportées révèlent une affinité préférentielle du domaine SRA pour l'ADN hémi-méthylé, il est important de noter que les différences d'affinité du domaine SRA observées pour les divers types d'ADN, c'est à dire non-méthylés, bi-méthylés ou hémi-méthylés, sont variables d'une étude à l'autre (Arita et al., 2008; Avvakumov et al., 2008; Bostick et al., 2007; Frauer et al., 2011; Greiner et al., en préparation; Hashimoto et al., 2008; Qian et al., 2008; Rottach et al., 2010 ). Ainsi, alors que trois études révèlent l'absence totale de liaison du domaine SRA à l'ADN non-méthylé (Arita et al., 2008; Bostick et al., 2007; Unoki et al., 2004), quatre autres indiquent une différence d'affinité du domaine SRA relativement faible entre l'ADN hémi-méthylé et non-méthylé (Frauer et al., 2011; Greiner et al., en préparation; Qian et al., 2008; Rottach et al., 2010 ). D'autre part, le domaine SRA pourrait se lier à l'ADN non-méthylé même en l'absence de son site de reconnaissance conventionnel CpG (Frauer et al., 2011; Greiner et al., en préparation). En effet, il n'est pas étonnant que le domaine SRA puisse se lier à tout type d'ADN puisque la protéine est chargée positivement à pH neutre, la liaison SRA/ADN présentant une part d'interactions électrostatiques. De plus, il serait surprenant qu'un simple groupement méthyle puisse conférer à la protéine des différences d'affinités très élevées entre les divers types d'ADN. Ainsi, dans le modèle que nous proposons, malgré sa préférence pour l'ADN hémi-méthylé, la protéine UHRF1 serait un opérateur de lecture capable de se lier à n'importe quel

ADN via son domaine SRA. Lors du processus de réplication, en se déplaçant sur l'ADN par diffusion, elle permettrait de repérer les sites hémi-méthylés et de basculer la cytosine méthylée pour permettre son ancrage et induire le recrutement de la Dnmt1 au niveau de la cytosine à méthyler en amont de la séquence d'ADN (Bronner et al., 2010; Greiner et al., en préparation; Hashimoto et al., 2009). *In vivo*, ce seraient donc le basculement de base induit par le domaine SRA, et la coopération de la Dnmt1, ainsi que les multiples interactions de UHRF1 avec les histones et les protéines de régulation de la chromatine qui favoriseraient l'ancrage de la protéine au niveau de sites stratégiques du génome. En effet, la liaison du domaine TTD à l'histone H3 étant nécessaire à la méthylation de maintenance de l'ADN (Rothbart et al., 2012), on peut imaginer que les différents domaines de UHRF1 se coordonnent pour une duplication fidèle des profils de méthylation de l'ADN.

En plus des 5-méthylcytosine (5mC), des études génomiques ont récemment révélé l'existence de 5-hydroxyméthylcytosines (5hmC) au sein des séquences d'ADN, résultant de l'oxydation de la 5mC catalysée par la famille de protéines Tet (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009). Les sites 5hmC joueraient un rôle dans le processus de déméthylation de l'ADN. D'autre part, ils présentent des profils particuliers et pourraient constituer une modification épigénétique à part entière. Une étude récente révèle que le domaine SRA de la protéine UHRF1 est capable de reconnaître la 5hmC et d'induire son basculement avec la même affinité que pour la 5mC (Frauer et al., 2011). Le rôle de ce mécanisme de reconnaissance des sites 5hmC par le domaine SRA de UHRF1 reste encore à déterminer. Cependant, il paraît évident qu'il pourrait jouer un rôle important dans la régulation de l'expression des gènes, d'une manière similaire à l'analogue 5mC, en permettant le recrutement de facteurs de régulation spécifiques au niveau de régions ciblées du génome.

Le domaine SRA de UHRF1 apparaît donc comme un acteur essentiel de la régulation épigénétique, permettant la connexion entre la méthylation de l'ADN et les modifications des histones.

## 2.5. Le domaine RING ("Really Interesting New Gene domain")

Au niveau de la région C-terminale de hUHRF1, on trouve un domaine en doigts de zinc de type RING composé des acides aminés 724 à 763. Ce domaine RING possède la seule

activité enzymatique connue de la protéine UHRF1, une activité E3 ligase. Cette enzyme catalyse la dernière étape d'une cascade de réactions permettant la liaison covalente de l'ubiquitine sur les histones ou d'autres protéines, leur mono-ubiquitinylation participant à des fonctions régulatrices alors que leur poly-ubiquitinylation entraîne leur dégradation par le protéasome. Ainsi, des études ont démontré que la protéine UHRF1 possède une forte activité d'auto-ubiquitinylation grâce à son domaine RING, ce processus constituant un mécanisme d'auto-régulation (Felle et al., 2011; Jenkins et al., 2005; Karagianni et al., 2008; Qin et al., 2011). D'autre part, grâce à son activité E3 ligase, le domaine RING serait également capable d'ubiquitinyler les histones avec une préférence marquée pour l'histone H3 nucléosomique (Citterio et al., 2004; Karagianni et al., 2008). En plus de l'histone H3, le domaine RING est susceptible d'induire l'ubiquitinylation d'autres protéines. Ainsi, en favorisant l'ubiquitinylation de la Dnmt1, le domaine RING de la protéine UHRF1 participe activement au processus de dégradation de cette enzyme (Bronner, 2011; Du et al., 2010; Felle et al., 2011; Qin et al., 2011). Une autre protéine ubiquitinylée grâce à l'activité E3 ligase de UHRF1 est la protéine suppresseur de tumeur PML ("promyélocytic leukemia"), impliquée dans les leucémies promyélocyaires et dans la tumorigenèse de multiples formes de cancers. En facilitant l'ubiquitinylation et la dégradation de la protéine PML, la surexpression de la protéine UHRF1 observée dans certains cancers favoriserait le développement de la tumeur (Guan et al., 2012). Enfin, le domaine RING permettrait le recrutement de protéines de régulation épigénétique telles que l'histone méthyltransférase G9a (Kim et al., 2009b).

### 3. Expression cellulaire de hUHRF1

Les profils d'expression des différents membres de la famille UHRF sont extrêmement variables selon les tissus et les espèces. Concernant la protéine hUHRF1, sa production est remarquablement corrélée avec la prolifération cellulaire. En effet, alors que dans les tissus hautement différenciés, tels que le système nerveux central et les leucocytes périphériques, les ARNm d'UHRF1 sont pratiquement absents, ils sont retrouvés de manière particulièrement abondante dans les cellules proliférantes et/ou non différenciées telles que dans les tissus fœtaux, le thymus et la moelle osseuse (Bronner et al., 2007a; Hopfner et al., 2000; Pichler et al., 2011).



La protéine hUHRF1 est exprimée dans le noyau. Dans les cellules saines, elle présente deux pics d'expression, à la fin de la phase G1 où elle se concentre au niveau de l'hétérochromatine (Bostick et al., 2007), et lors de la transition G2/M (Jeanblanc et al., 2005; Mousli et al., 2003). A l'inverse, dans les cellules cancéreuses, l'expression de la protéine est indépendante du cycle cellulaire et se retrouve de manière constante à des taux relativement élevés (Jeanblanc et al., 2005; Mousli et al., 2003). Ainsi, une surexpression de hUHRF1, souvent associée à la progression de la maladie, a été observée dans divers cancers tels que le cancer du sein (Hopfner et al., 2000; Jenkins et al., 2005; Mousli et al., 2003; Unoki et al., 2004), de la prostate (Babbio et al., 2012; Jenkins et al., 2005), de la vessie (Unoki et al., 2009b; Yang et al., 2012) et des poumons (Daskalos et al., 2011; Hopfner et al., 2000; Unoki et al., 2010), le cancer colorectal (Sabatino et al., 2012; Wang et al., 2012), les lésions cervicales (Lorenzato et al., 2005), les gliomes (Oba-Shinjo et al., 2005), le rhabdomyosarcome (Schaaf et al., 2005) et l'adénocarcinome pancréatique (Crnogorac-Jurcevic et al., 2005).

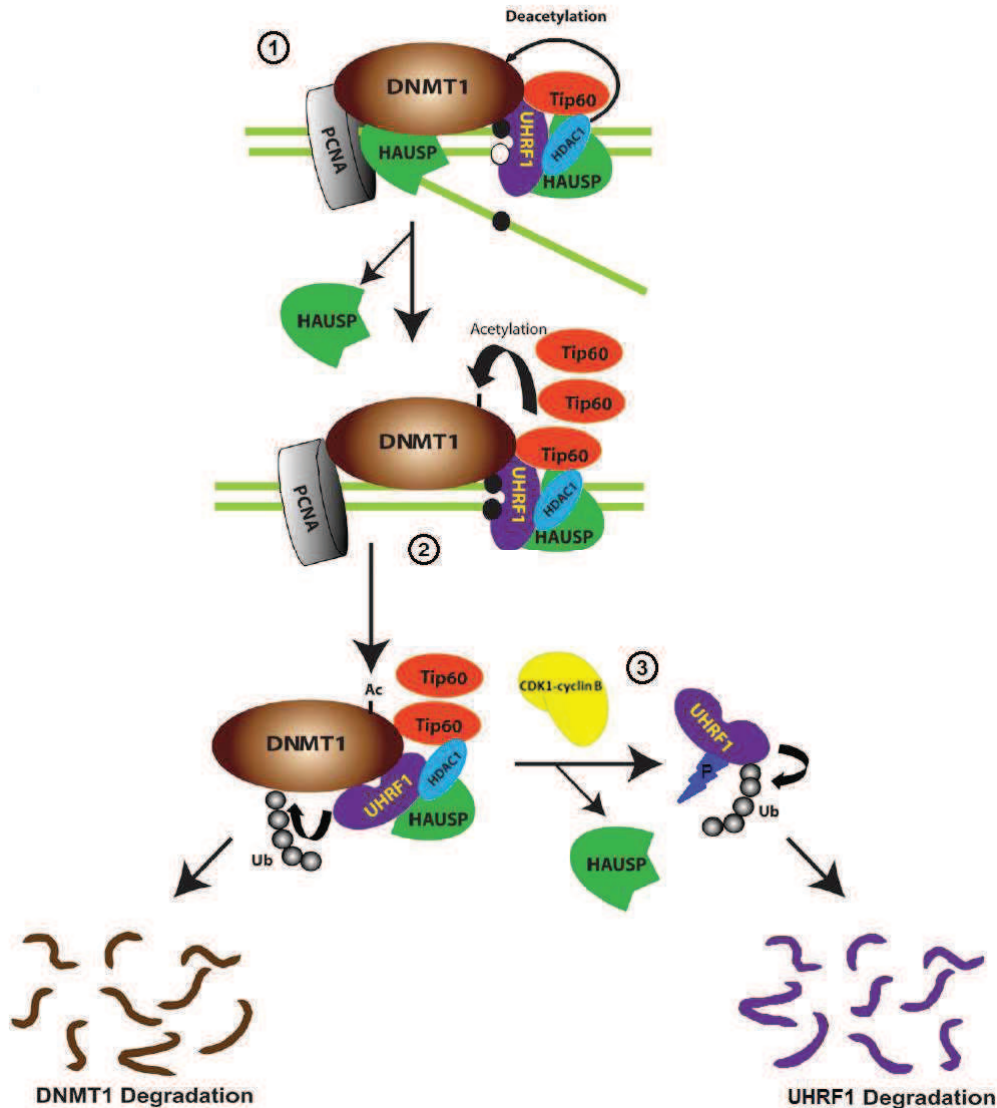
#### 4. Les différents rôles de la protéine hUHRF1

##### 4.1. Auto-régulation de UHRF1 et régulation de la Dnmt1

Grâce à l'activité E3 ligase de son domaine RING, la protéine UHRF1 est impliquée dans sa propre régulation et celle de la Dnmt1. En effet, la protéine UHRF1 possède une forte activité d'auto-ubiquitinylation (Felle et al., 2011; Jenkins et al., 2005; Karagianni et al., 2008; Qin et al., 2011). D'autre part, l'acétylation de la Dnmt1 par l'acétyltransférase Tip60 favorise l'ubiquitinylation de la Dnmt1 par UHRF1 (Bronner, 2011; Du et al., 2010). Les deux protéines, UHRF1 et Dnmt1, sont alors ciblées et dégradées par le protéasome. A l'inverse, l'enzyme HAUSP ("herpes virus-associated ubiquitin specific protease" ou Usp7 pour "ubiquitin specific peptidase 7"), qui interagit de manière directe avec les deux protéines UHRF1 et Dnmt1, permet leur dé-ubiquitinylation et les protège de la dégradation par le protéasome (Bronner, 2011; Du et al., 2010; Felle et al., 2011; Ma et al., 2012; Qin et al., 2011). Ainsi, UHRF1 et HAUSP ont des effets opposés et jouent un rôle dans la stabilisation



des protéines UHRF1 et Dnmt1 en régulant leur statut d'ubiquitinylation de manière dépendante du cycle cellulaire.



*Figure 19 : Modèle de régulation des protéines UHRF1 et Dnmt1 (D'après (Du et al., 2010) et (Ma et al., 2012))*

1) HAUSP interagit avec Dnmt1 et UHRF1 au niveau de la chromatine, dans un complexe contenant Tip60, HDAC1 et PCNA. HDAC1 prévient l'acétylation de la Dnmt1 et HAUSP empêche UHRF1 de s'auto-ubiquitinyler et d'ubiquitinyler la Dnmt1, prévenant ainsi leur dégradation par le protéasome.

2) Lors de la phase S, après la réplication des profils de méthylation de l'ADN médiée par la Dnmt1 en association avec UHRF1, HAUSP se dissocie de la Dnmt1 qui n'est donc plus protégée. L'augmentation de Tip60 induit alors l'acétylation de la Dnmt1 qui à son tour entraîne l'ubiquitinylation de la Dnmt1 par la protéine UHRF1.

3) Lors de la phase M, le complexe cycline B-Cdk1 permettrait la phosphorylation de la protéine UHRF1, entraînant la dissociation d'HAUSP. UHRF1 pourrait alors s'auto-ubiquitinyler induisant sa dégradation par le protéasome.

Dans le modèle actuellement proposé (Figure 19) (Bronner, 2011; Du et al., 2010; Felle et al., 2011; Ma et al., 2012) HAUSP interagit avec Dnmt1, ce duplex étant recruté par UHRF1 au niveau de la chromatine, dans un complexe contenant Tip60, HDAC1 et PCNA. HDAC1 prévient l'acétylation de la Dnmt1 et HAUSP empêche UHRF1 de s'auto-ubiquitinyler et d'ubiquitinyler la Dnmt1, prévenant ainsi leur dégradation par le protéasome. Outre son rôle de protection, HAUSP stimulerait l'activité enzymatique de la Dnmt1 (Felle et al., 2011). Lors de la phase S, après la répllication des profils de méthylation de l'ADN médiée par la Dnmt1 en association avec UHRF1, HAUSP se dissocie de la Dnmt1 qui n'est donc plus protégée. L'augmentation de Tip60 induit alors l'acétylation de la Dnmt1 qui à son tour entraîne l'ubiquitinylation de la Dnmt1 par la protéine UHRF1. Lors de la phase M, le complexe cycline B-Cdk1 permettrait la phosphorylation de la protéine UHRF1, entraînant la dissociation d'HAUSP (Ma et al., 2012). UHRF1 pourrait alors s'auto-ubiquitinyler induisant sa dégradation par le protéasome.

#### 4.2. hUHRF1 dans la régulation du cycle cellulaire

Plusieurs études révèlent l'importance de la protéine UHRF1 dans la progression du cycle cellulaire. Ainsi, la suppression de la protéine hUHRF1 dans divers types cellulaires entraîne l'inhibition de leur progression avec un arrêt des cellules en phase G1 ou G2/M pouvant parfois entraîner l'apoptose (Abbadly et al., 2005; Bronner et al., 2007a; Jenkins et al., 2005; Tien et al., 2011; Unoki et al., 2004). D'autre part, la surexpression de hUHRF1 augmente la proportion de cellules dans les phases S et G2/M (Jeanblanc et al., 2005). L'induction de diverses voies de signalisation cellulaire telles que la voie de contrôle p53/p21Cip1/WAF1 activée en réponse aux dommages de l'ADN (Arima et al., 2004; Bronner et al., 2007a; Jenkins et al., 2005), la voie du TCR (récepteur des cellules T) (Abbadly et al., 2005), celle de p73 (Abusnina et al., 2011; Alhosin et al., 2010), ou encore l'inhibition de celle de ERK1/2 (Fang et al., 2009), est associée à une réduction de l'expression de la protéine hUHRF1 et provoque également l'arrêt du cycle cellulaire en G1 pouvant entraîner la mort cellulaire (Figure 20). La protéine hUHRF1 semble donc posséder une propriété anti-apoptotique et joue un rôle essentiel dans l'entrée en phase S du cycle cellulaire et lors de la transition G2/M. Bien que de nombreuses études révèlent une corrélation entre la réduction de

l'expression de UHRF1 et l'arrêt du cycle cellulaire ou l'apoptose, les mécanismes moléculaires de ce processus restent encore indéterminés.

L'entrée et la progression des cellules dans le cycle cellulaire sont régulées par l'activation séquentielle de kinases cycline-dépendantes (Cdk) en complexe avec des cyclines. Lorsqu'ils sont activés lors de la prolifération cellulaire, ces complexes vont induire l'hyperphosphorylation de la protéine du rétinoblastome (pRb) qui va alors se dissocier du facteur de transcription E2F-1 (Figure 20). Ce facteur de transcription est un acteur clé de la transition G1/S, et une fois libéré de son inhibition par pRb, il va permettre l'expression de diverses protéines impliquées dans la régulation du cycle cellulaire telles que des cyclines, des Cdk, des régulateurs des points de contrôle du cycle cellulaire, ou encore des protéines de réparation de l'ADN. Le promoteur du gène codant pour *hUHRF1* contenant trois sites de liaison pour E2F-1, le rôle de la protéine hUHRF1 dans la transition du cycle cellulaire semble être contrôlé par cette famille de facteurs de transcription (Arima et al., 2004; Mousli et al., 2003; Unoki et al., 2004). Dans la progression du cycle cellulaire, E2F1 une fois dissocié de pRb permet donc d'activer l'expression de la protéine hUHRF1. D'autre part, la protéine hUHRF1 peut réguler négativement pRb en empêchant sa transcription par liaison au promoteur de son gène *RBI* méthylé, ou en interagissant directement avec la protéine pRb hyperphosphorylée (Jeanblanc et al., 2005). En absence de pRb, le facteur E2F1 peut alors activer les gènes nécessaires à l'entrée en phase S du cycle cellulaire. En régulant négativement pRb par ces deux mécanismes, la protéine hUHRF1 favorise donc la transition G1/S du cycle cellulaire. D'autre part, lors de la transition G1/S, la phosphorylation d'hUHRF1 induite par la PKA (protéine kinase A) via la voie de signalisation cAMP (adénosine monophosphate cyclique) augmenterait la liaison d'hUHRF1 au promoteur ICB2 du gène topoisomérase II $\alpha$ , favorisant son expression (Trotzler et al., 2004). La topoisomérase II $\alpha$  étant un acteur essentiel de la réplication, son activation par la protéine UHRF1 favoriserait également l'entrée en phase S du cycle cellulaire.

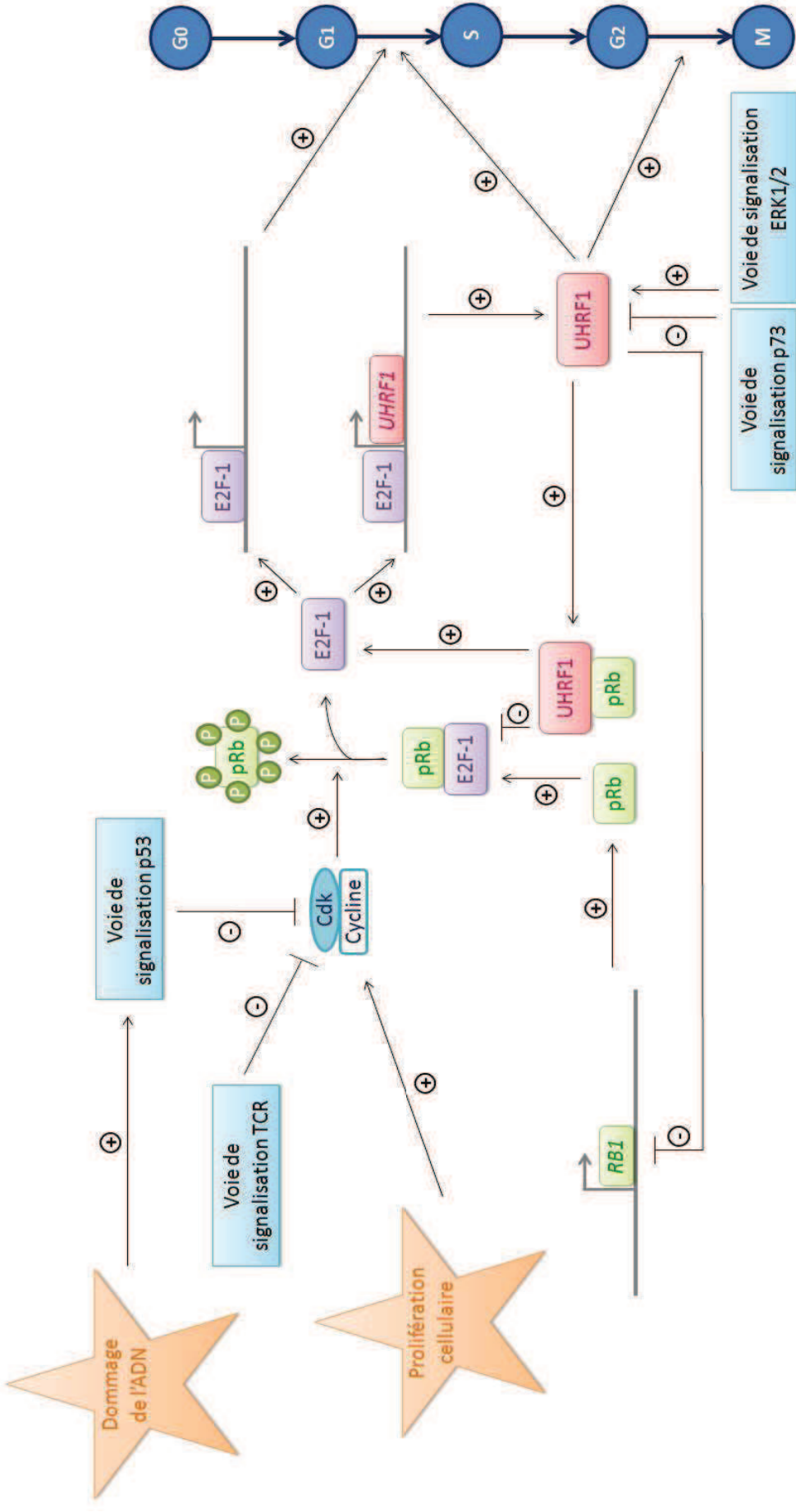


Figure 20 : Implication de UHRF1 dans la régulation du cycle cellulaire, au cours de la prolifération cellulaire, en réponse aux dommages de l'ADN, ou après induction de diverses voies de signalisation. Cdk : kinases cycline-dépendantes ; pRb : protéine du rétinoblastome codée par le gène RB1 ; E2F-1 : facteur de transcription contrôlant l'expression de protéines de régulation du cycle cellulaire.

Outre son rôle dans la régulation du cycle cellulaire, la protéine UHRF1 semble jouer un rôle dans la réponse aux dommages de l'ADN. En effet, dans différents types cellulaires, on observe une diminution significative de la protéine hUHRF1 et de son ARNm lorsque l'ADN est endommagé (Arima et al., 2004; Bronner et al., 2007a; Jenkins et al., 2005). La réduction constatée de hUHRF1 semble être provoquée à la fois par une réduction de sa transcription et par la stimulation d'une voie protéolytique (Arima et al., 2004). Dans ce contexte, l'expression de la protéine hUHRF1 est régulée par la voie de contrôle p53/p21Cip1/WAF1 activée en réponse aux dommages de l'ADN (Figure 20). Le complexe p21Cip1/WAF1 inhibe l'activité des Cdk et empêche la phosphorylation de pRb qui reste lié au facteur E2F1. L'expression de la protéine hUHRF1 est inhibée et le cycle cellulaire bloqué en G1/S. Il est important de noter que même dans les cellules cancéreuses dans lesquelles p53 est inactif, la réduction de l'expression de hUHRF1 peut seule empêcher la prolifération cellulaire et induire l'apoptose (Abbady et al., 2003). Ces observations suggèrent que la protéine hUHRF1 constitue un élément de la voie de réponse aux dommages de l'ADN. Bien que présentant des résultats quelque peu contradictoires avec cette première série de travaux, une étude récente vient confirmer l'implication directe de la protéine UHRF1 dans ce processus (Tien et al., 2011). En effet, dans cette étude, la protéine hUHRF1 est retrouvée accumulée au niveau des sites endommagés de l'ADN. D'autre part, l'inhibition de la protéine hUHRF1 dans des cellules cancéreuses induit directement l'activation de la voie de réponse aux dommages de l'ADN indépendamment de p53 et grâce aux protéines kinases ATM (ataxia telangectasia mutated) et/ou ATR (ATM and Rad3-related) qui inhibent la phosphorylation de la Cdk1, induisant l'arrêt du cycle cellulaire en G2/M et l'apoptose. Ces résultats suggèrent que la protéine UHRF1 joue un rôle important dans la réponse aux dommages de l'ADN et pourrait être associée au système de réparation de l'ADN. Chez les souris, on observe ainsi mUHRF1 en complexe avec des facteurs de réparation tels que PARP-1, XRCC1, TopBP1, PCNA et Emel (Mistry et al., 2008; Unoki et al., 2004), et les cellules dépourvues en mUHRF1 sont plus sensibles aux agents causant des dommages à l'ADN (Muto et al., 2002).

### 4.3. hUHRF1 dans la régulation épigénétique

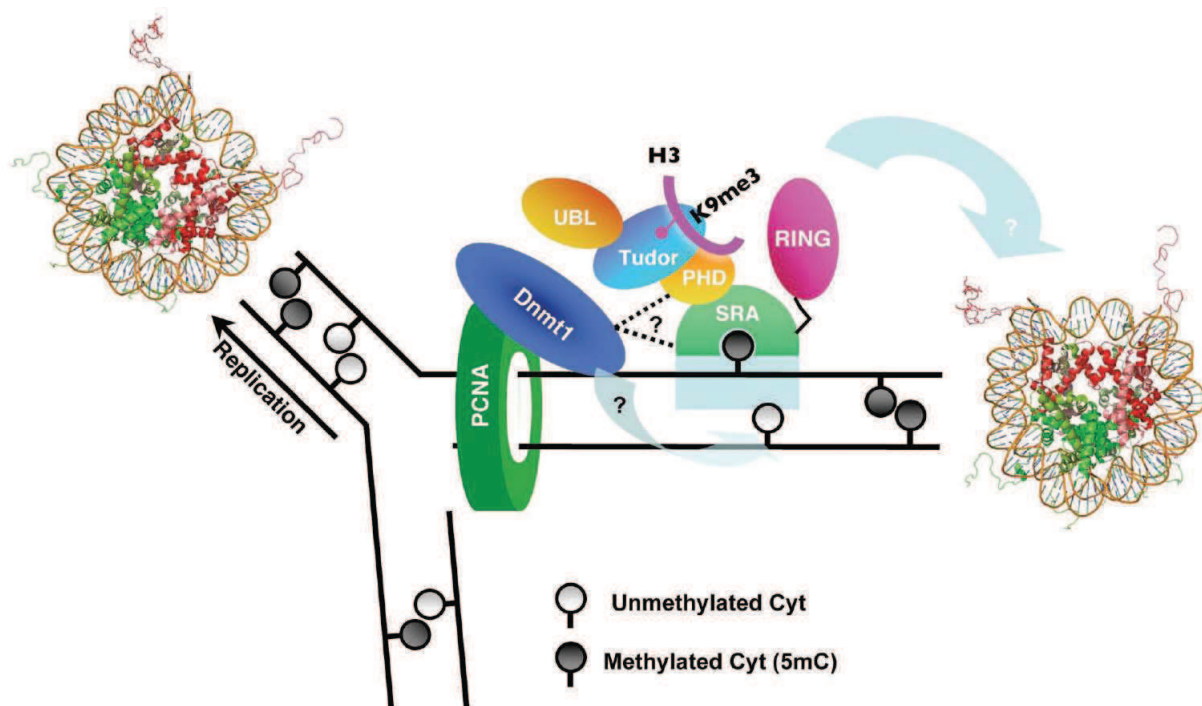
De nombreuses études ont révélé l'importance de la protéine UHRF1 dans la régulation et la maintenance du code épigénétique (Bronner et al., 2010; Hashimoto et al., 2009; Hashimoto et al., 2010). En recrutant de nombreux partenaires de régulation au niveau de la chromatine, la protéine UHRF1 constituerait un point focal, et permettrait la formation d'un gros complexe macromoléculaire appelé ECREM ("Epigenetic Code REplication Machinery") responsable de la transmission du code épigénétique (Bronner et al., 2007b; Kim et al., 2009b; Sharif et al., 2007). Dans ce complexe, la protéine UHRF1 jouerait un rôle central puisqu'elle est la seule protéine connue capable de coordonner directement la méthylation de l'ADN et les modifications des histones. D'autre part, en étant capable de lire une marque épigénétique et de recruter le partenaire capable de catalyser la synthèse de cette même marque, la protéine UHRF1 serait un acteur clé de la réplication du code épigénétique lors des divisions cellulaires.

D'une part, essentiellement grâce à son domaine SRA, la protéine UHRF1 joue un rôle fondamental dans la transmission des profils de méthylation de l'ADN. Ainsi, de la même manière que les mutants Dnmt1, les souris "knockout" pour UHRF1 meurent durant l'embryogenèse, et la déplétion de hUHRF1 ou mUHRF1 entraîne une perte massive de la méthylation globale et locale de l'ADN (Bostick et al., 2007; Sharif et al., 2007). La protéine UHRF1 favorise la localisation de la Dnmt1 au niveau de la chromatine, préférentiellement au niveau des sites hémi-méthylés, et s'associe avec la Dnmt1 et PCNA au niveau de la fourche de réplication (Bostick et al., 2007; Sharif et al., 2007). En guidant la Dnmt1 au niveau de l'ADN hémi-méthylé, la protéine UHRF1 constituerait donc un élément essentiel dans la maintenance de la méthylation de l'ADN et la réplication de l'hétérochromatine (Bostick et al., 2007; Karagianni et al., 2008; Papait et al., 2008; Sharif et al., 2007).

D'autre part, grâce à ses domaines TTD et PHD, la protéine UHRF1 est capable de lire le code histone et serait susceptible de le transmettre en recrutant les enzymes de modification de la chromatine correspondantes. Ces deux domaines semblent fonctionner en association pour reconnaître spécifiquement et simultanément les marques H3K4me0, H3R2me0 et H3K9me3 (Arita et al., 2012; Nady et al., 2011; Xie et al., 2012). En reconnaissant ces modifications et en recrutant G9a et HDAC1 au niveau de ces régions spécifiques du génome, UHRF1 favoriserait la formation de l'hétérochromatine et sa propagation. D'autre part, lors de



la réplication, la protéine UHRF1 serait capable de reconnaître H3K9me3 sur l'histone parentale et de recruter la méthyltransférase G9a qui catalyserait spécifiquement la méthylation d'H3K9 sur l'histone nouvellement assemblée (Bronner et al., 2010; Hashimoto et al., 2010). Grâce à ce mécanisme, la protéine UHRF1 serait donc capable de répliquer le code histone. Enfin, grâce à son domaine RING, la protéine UHRF1 joue également un rôle d'effecteur en favorisant l'ubiquitinylation des histones (Citterio et al., 2004; Karagianni et al., 2008), entraînant l'activation ou la répression de gènes.



*Figure 21 : Modèle proposé pour la réplication simultanée des profils de méthylation de l'ADN et de modifications des histones induite par UHRF1 (Hashimoto et al., 2009). L'existence simultanée de modules de reconnaissance de l'ADN méthylé (via SRA) et des histones (via TTD et PHD) dans la protéine UHRF1 permettrait la maintenance et la conversion des marques épigénétiques de répression à la fois sur l'ADN et les histones.*

Grâce à ses différents domaines, la protéine UHRF1 serait donc capable de lire simultanément les modifications de l'ADN et des histones, d'intégrer ces informations, et d'induire en conséquence des modifications de la chromatine via son domaine RING ou en recrutant divers acteurs de la régulation épigénétique (Figure 21). Soutenant cette hypothèse, la protéine UHRF1 est particulièrement enrichie au niveau des nucléosomes lorsqu'ils contiennent H3K4me0/K9me3 et semble favoriser la méthylation de l'ADN au niveau de ces



régions (Bartke et al., 2010 ; Rothbart et al., 2012). Une étude récente révèle ainsi les modifications simultanées des profils de méthylation d'H3K9, d'acétylation et de méthylation de l'ADN dans des cellules cancéreuses dans lesquelles UHRF1 a été supprimé (Babbio et al., 2012). D'autre part, confirmant la liaison entre la modification des histones et la méthylation de l'ADN induite par UHRF1, la capacité de la protéine à réprimer l'expression de gènes cibles en favorisant la méthylation de leur promoteur est corrélée à la liaison d'UHRF1 au niveau des nucléosomes non modifiés sur H3R2 (Rajakumara et al., 2011). La protéine UHRF1 semble permettre une action synergique de la méthylation de l'ADN et des modifications des histones qui travaillent ensemble pour influencer la structure de la chromatine et l'expression des gènes. Cette synergie est particulièrement évidente dans la répression des gènes où UHRF1 favorise l'organisation de l'hétérochromatine en coordonnant la formation d'un complexe macromoléculaire capable de catalyser l'acétylation des histones via HDAC1, la méthylation d'H3K9 via G9a, et la méthylation de l'ADN via la Dnmt1 (Bronner et al., 2010).

Ainsi, grâce à son domaine SRA et ses domaines TTD et PHD, la protéine UHRF1 agit respectivement à l'interface entre l'ADN et les histones. Lors de la réplication, en favorisant simultanément la transmission des profils de méthylation de l'ADN via la Dnmt1 et celle des modifications des histones via G9a et HDAC1, elle jouerait un rôle essentiel dans la réplication du code épigénétique (Figure 21).

#### 4.4. hUHRF1 dans la cancérogenèse

La protéine hUHRF1 est surexprimée dans de nombreux cancers, probablement du fait de la dérégulation de divers processus dans les cellules cancéreuses, tels que la progression rapide des cellules et la fréquence des mécanismes de réparation de l'ADN. D'autre part, le gène *p53* est muté dans 50 % des cancers, son expression ou sa fonction étant supprimée (Hussain and Harris, 2000). La voie de *p53* régulant négativement l'expression de UHRF1 (Arima et al., 2004) par désactivation du facteur de transcription E2F-1, sa dérégulation pourrait expliquer la surexpression de la protéine UHRF1 observée dans certains cancers (Unoki et al., 2009a). Cette surexpression serait responsable du dérèglement de divers processus cellulaires. Ainsi, l'expression de la protéine UHRF1 dans les tissus tumoraux est corrélée à l'hyperméthylation de promoteurs de gènes suppresseurs de tumeurs sur lesquels la

protéine est retrouvée localisée et coordonnerait la répression des gènes (Unoki et al., 2004). Confirmant le rôle direct de UHRF1 dans l'établissement et/ou la transmission de ces répressions, la déplétion de la protéine dans les tissus concernés entraîne la réactivation des gènes suppresseurs de tumeur, accompagnée d'une déméthylation de l'ADN et de l'introduction de modifications activatrices des histones (Babbio et al., 2012; Daskalos et al., 2011; Sabatino et al., 2012; Wang et al., 2012). D'autre part, UHRF1 jouant un rôle dans la régulation du cycle cellulaire, sa surexpression favoriserait le maintien des cellules cancéreuses dans un état prolifératif empêchant leur différenciation. En réprimant de manière permanente l'expression de gènes impliqués dans la progression du cycle cellulaire, tels que *RBI* et *p16<sup>INK4A</sup>*, la protéine favoriserait ainsi le développement des tumeurs (Achour et al., 2008; Bronner et al., 2007a; Jeanblanc et al., 2005 ; Wang et al., 2012). Confirmant le rôle d'UHRF1 dans la progression des cancers, la déplétion de la protéine dans les tissus et les cellules cancéreuses réduit leur prolifération, diminue leur propriété de migration, et induit leur apoptose (Daskalos et al., 2011; Unoki et al., 2004; Wang et al., 2012).

Ainsi, en entraînant la répression de nombreux gènes suppresseurs de tumeur, en favorisant la prolifération cellulaire aberrante, la migration, et l'échappement à l'apoptose des cellules cancéreuses, la protéine UHRF1 joue un rôle central dans la cancérogenèse et/ou dans le maintien du phénotype cancéreux.

## 5. UHRF1 comme marqueur diagnostique et pronostique du cancer

La surexpression de la protéine UHRF1 observée dans les cancers est souvent corrélée à l'évolution de la maladie. Ainsi, dans le cancer de la vessie, elle est associée à un risque augmenté de la progression après résection transurétrale (Unoki et al., 2009b), et dans les cancers non-invasifs elle est liée au stade de la tumeur et à la récurrence de la maladie, avec un temps de survie moyen diminué lorsque l'expression d'UHRF1 est élevée (Yang et al., 2012). Dans le cancer des poumons, l'expression d'UHRF1 est associée au type histologique de la tumeur (adénocarcinomes ou non), à la propagation de la maladie cancéreuse sur le site de la tumeur primitive (facteur T) et dans les ganglions lymphatiques voisins (facteur N), et à un pauvre pronostic chez les patients atteints d'un cancer bronchique non-lié à de petites cellules (Unoki et al., 2010). L'expression de la protéine UHRF1 est également liée au stade de la tumeur dans les cancers colorectaux (Sabatino et al., 2012; Wang et al., 2012) et au taux

de survie après prostatectomie dans le cancer de la prostate (Babbio et al., 2012). L'expression de la protéine UHRF1 étant facilement détectable par immunohistochimie, elle apparaît donc comme un nouveau marqueur diagnostique et pronostique potentiel dans un certains nombre de cancers (Babbio et al., 2012; Crnogorac-Jurcevic et al., 2005; Unoki et al., 2010; Unoki et al., 2009b; Wang et al., 2012; Yang et al., 2012).

## 6. UHRF1 comme cible thérapeutique potentielle

Plusieurs raisons laissent à penser que la protéine UHRF1 pourrait être une cible thérapeutique intéressante dans les traitements anticancéreux. En effet, la protéine UHRF1, fréquemment surexprimée dans les cancers, est impliquée dans la cancérogenèse en raison de son rôle dans la progression du cycle cellulaire et de ses propriétés anti-apoptotiques. Le fait que la déplétion d'UHRF1 dans les cellules cancéreuses puisse réduire la prolifération et la croissance cellulaire indépendante de l'ancrage, réactiver des gènes suppresseurs de tumeur, et entraîner l'apoptose, indique que l'utilisation d'agents thérapeutiques ciblant la protéine pourrait être bénéfique dans le traitement des cancers (Daskalos et al., 2011; Wang et al., 2012). De plus, la protéine UHRF1 ne semblant pas absolument requise pour la prolifération cellulaire des cellules saines, son inhibition permettrait d'affecter la prolifération des cellules cancéreuses sans affecter la viabilité des cellules souches embryonnaires et des cellules saines en générale (Bronner et al., 2007a; Muto et al., 2002). D'autre part, des études révèlent que la diminution de l'expression de UHRF1 augmente la sensibilité des cellules cancéreuses à l'hydroxyurée ou d'autres agents anticancéreux (Bronner et al., 2007a; Jenkins et al., 2005; Un et al., 2006). La diminution de l'expression de la protéine UHRF1 permettrait d'augmenter la chimio-sensibilité des cellules cancéreuses, probablement grâce à la réexpression des gènes suppresseurs de tumeur, et pourrait donc constituer une stratégie intéressante pour accompagner les traitements anticancéreux déjà existants (Bronner et al., 2007a).

## II- OBJECTIFS

On découvre aujourd'hui que les altérations épigénétiques sont impliquées dans le développement de nombreuses maladies et particulièrement dans le cancer (Portela and Esteller, 2010). L'étude au niveau fondamental des mécanismes régissant la régulation épigénétique apparaît donc comme une stratégie intéressante pour découvrir de nouveaux agents thérapeutiques agissant par des mécanismes originaux. Parmi les acteurs de la régulation épigénétique, la protéine UHRF1 possède un ensemble de caractéristiques lui permettant de maintenir et de transmettre des modifications épigénétiques grâce aux différents domaines qui la constituent (Bronner et al., 2010; Hashimoto et al., 2009; Hashimoto et al., 2010). Elle apparaît donc comme une cible de choix pour le développement de nouvelles molécules qui pourraient constituer une alternative aux inhibiteurs des Dnmts utilisés dans les traitements anticancéreux. Par ailleurs, l'étude de cette protéine apporte un nouveau niveau de compréhension dans la régulation des mécanismes épigénétiques.

Dans le processus de duplication des profils de méthylation de l'ADN, la protéine UHRF1 permettrait le recrutement de la Dnmt1 au niveau des sites CpG hémi-méthylés reconnus par son domaine SRA (Bostick et al., 2007; Sharif et al., 2007). Ce domaine se comporte comme une main dont les doigts font basculer la cytosine méthylée dans le grand sillon de l'ADN (Arita et al., 2008; Avvakumov et al., 2008; Hashimoto et al., 2008). Ce mécanisme de reconnaissance des sites hémi-méthylés, lui permettrait de s'ancrer à des endroits stratégiques, facilitant le recrutement de la Dnmt1. Dans ce contexte, notre projet vise à comprendre les mécanismes intimes de l'interaction du domaine SRA de UHRF1 avec l'ADN hémi-méthylé et du basculement de la cytosine méthylée. Pour répondre à ces questions, la 2-aminopurine (2-Ap) nous est apparue comme un outil de choix. En effet, la 2-Ap est un analogue fluorescent de l'adénine pouvant substituer une base naturelle dans un ADN double brin et y jouer le rôle de rapporteur fluorescent. La 2-Ap est sensible à son microenvironnement et plus particulièrement à l'empilement ("stacking") avec ses bases voisines. Pour cette raison, elle donne des informations locales au sein d'un acide nucléique. Dans nos expériences, l'introduction de la 2-Ap à différentes positions dans des oligonucléotides (ODNs) de 12 pb présentant un unique site CpG hémi-méthylé nous a permis d'élucider les changements de conformation locaux induits par la liaison du domaine SRA au voisinage de son site de liaison. D'autre part, la comparaison de l'effet de la liaison du

domaine SRA sur la dynamique des acides nucléiques dans des ODNs dont le site de reconnaissance CpG est hémi-méthylé, non-méthylé ou bi-méthylé nous a permis de caractériser plus avant la liaison du domaine SRA à l'ADN. Grâce à de multiples techniques physiques telles que la spectroscopie de fluorescence à l'état stationnaire, la fluorescence résolue en temps, et le transfert d'énergie de fluorescence (FRET), nous avons ainsi approfondi nos connaissances au niveau moléculaire des mécanismes d'interaction du domaine SRA de UHRF1 avec l'ADN.

### III- MATÉRIELS ET MÉTHODES

#### A. Matériels

##### 1. Domaine SRA (408-643) de la protéine UHRF1

GHM<sup>408</sup>KEC TIVPS NHYGP IPGIP VGTMW RFRVQ VSESG VHRPH VAGIH GRSND GAYSL VLAGG  
YEDDV DHGNF FTYTG SGRD LSGNK RTAEQ SCDQK LTNTN RALAL NCFAP INDQE GAEAK  
DWRSK KPVRV VRNVK GGKNS KYAPA EGNRY DGIYK VVKYW PEKSK SGFLV WRYLL RRDDD  
EPGPW TKEGK DRIKK LGLTM QYPEG YLEAL ANRER EKENS KREEE EQQEG GFASP RTG<sup>643</sup>

*Figure 22* : Séquence d'acides aminés après digestion à la rTEV.  $M = 26.749 \text{ Da}$  ;  $\epsilon = 43.890 \text{ M}^{-1}\text{cm}^{-1}$  ; Nombre d'acides aminés = 239.

##### 1.1. Production et purification de la protéine

Le domaine SRA de hUHRF1 (résidus 408 à 643) est produit et purifié d'après le protocole de Delagoutte et al. 2008 (Delagoutte et al., 2008) par Catherine Birck grâce à une collaboration avec le Département de Biologie Structurale Intégrative de l'IGBMC. Le plasmide utilisé contient la séquence codant pour le domaine SRA (408-643) intégrée dans un vecteur d'expression pHGWA qui permet de produire une protéine avec une étiquette hexahistidine en N-terminal ainsi qu'un site de clivage à la rTEV (recombinante tobacco etch virus). Après digestion à la rTEV, trois résidus (GHM) sont conservés en N-terminal de la protéine SRA (Figure 22).

Des bactéries *Escherichia coli* BL21(DE3) transformées par le plasmide d'expression de SRA et le plasmide pLysS, qui permet de bloquer la production de la protéine hors induction par l'IPTG (isopropyl  $\beta$ -D-1-thiogalactopyranoside), sont mises en culture dans du milieu LB contenant  $100 \mu\text{g.mL}^{-1}$  d'ampicilline et  $35 \mu\text{g.mL}^{-1}$  de chloramphénicol et incubées à  $37^\circ\text{C}$  jusqu'à l'obtention d'une  $\text{DO}_{600}$  de  $\sim 0,4$ . La température est réduite à  $20^\circ\text{C}$  pendant 30 min et, afin d'induire l'expression de la protéine, de l'IPTG est ajouté à une concentration finale de 0,5 mM. La croissance des bactéries est poursuivie pendant 5 h après

lesquelles les bactéries sont récoltées par centrifugation et lysées par sonication dans du tampon A (30 mM Tris-HCl pH 7, 150 mM NaCl, 10 mM imidazole, 200 U de benzonase, 5 mM  $\beta$ -mercaptoéthanol et 1 mM PMSF) à 4°C. Le lysat est clarifié par centrifugation à 14.000 rpm pendant 45 min à 4°C puis chargé sur une colonne d'affinité au nickel (5 mL Ni<sup>2+</sup>-loaded HiTrap Chelating HP column, GE Healthcare). La protéine SRA est éluée avec un gradient linéaire d'imidazole de 10 à 300 mM dans le tampon A. Les fractions contenant la protéine SRA sont identifiées par gel SDS-PAGE, concentrées, et dialysées pendant la nuit contre le tampon A. La protéine dialysée est digérée par la protéase rTEV pendant 72 h à 4°C au ratio de 1 mg de protéase pour 40 mg de protéines puis le mélange réactionnel est chargé sur une colonne d'affinité au nickel. La protéine SRA retrouvée dans les fractions qui s'écoulent de la colonne est concentrée et chargée sur une colonne de chromatographie d'exclusion stérique (Superdex 75 HR 16/50, Amersham Biosciences) équilibrée avec 25 mM Tris-HCl pH 7,5, 50 mM NaCl, 0,5 mM PMSF, 2 mM TCEP (Tris(2-carboxyéthyl)phosphine). Les fractions contenant la protéine d'intérêt sont identifiées par gel SDS-PAGE et la protéine est concentrée sur des filtres à centrifuger (Amicon15 MWCO 10 kDa) à une concentration d'environ 2 mg/mL (65  $\mu$ M).

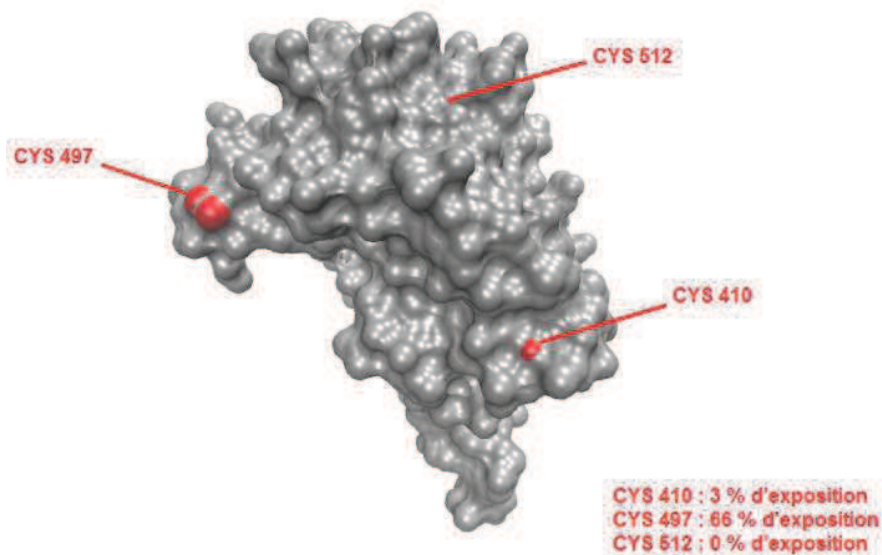
## 1.2. Stockage de la protéine

La protéine SRA ayant tendance à précipiter à la congélation, 5 % de glycérol sont ajoutés à la solution de protéine avant de l'aliqoter (150  $\mu$ L) et de la congeler rapidement dans de l'azote liquide afin d'éviter son agrégation. La protéine est ensuite conservée à -80°C.

## 1.3. Marquage de la protéine au Cy3

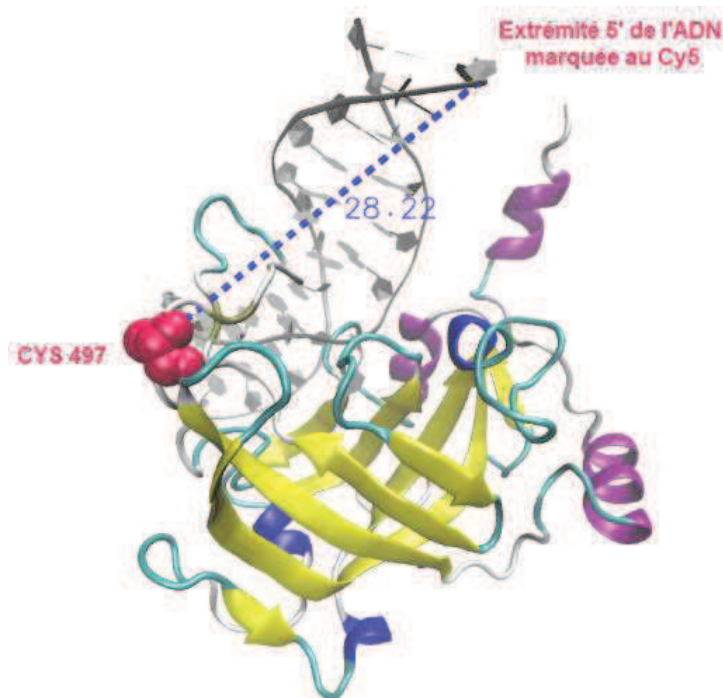
Afin de déterminer l'affinité de la protéine SRA (408-643) pour l'ADN par la technique de transfert d'énergie non-radiatif (FRET, Fluorescence Resonance Energy Transfer) (voir p. 94), la protéine est marquée sélectivement sur une cystéine par un fluorophore, le Cy3, et les oligonucléotides partenaires à leur extrémité 5' par du Cy5.





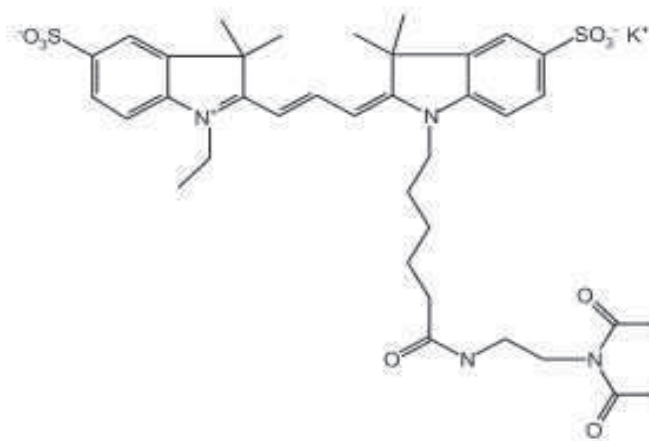
*Figure 23 : Degrés d'exposition des cystéines de la protéine SRA (408-643). Calculs effectués sur "<http://curie.utmb.edu/getarea.html>" et représentation sur VMD (code PDB 2PB7).*

La protéine SRA (408-643) possède trois cystéines (Figure 22) dont une seulement est exposée au solvant (Figure 23), la cystéine 497. Le marquage sélectif de cette protéine sur une cystéine est donc possible. De plus, la distance théorique entre le fluorophore de l'ADN en 5' et le fluorophore de la protéine sur la cystéine 497 est adéquate pour observer du transfert d'énergie entre les deux fluorophores lors d'une interaction ADN/protéine (Figure 24).



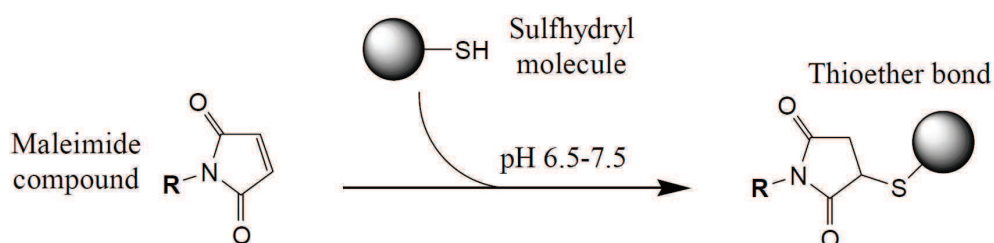
*Figure 24 : Représentation tridimensionnelle du complexe SRA/ADN. La distance entre la cystéine 497 de la protéine SRA et l'extrémité 5' de l'ADN est représentée par la ligne pointillée bleue (28,22 Å). Représentation sur VMD, code PDB 3CLZ.*

Pour effectuer le marquage de la protéine SRA (408-643), on utilise le kit de marquage "Amersham Cy3 Maleimide Monoreactive dye 5-pack". La sonde Cy3 fournie est une maléimide mono-fonctionnelle (Figure 25) permettant le marquage de composés contenant un groupement sulfhydryle libre tel que la cystéine.



*Figure 25 : Structure chimique du Cy3.*

La monodose de poudre de sonde Cy3 ( $M = 765,95 \text{ g. mol}^{-1}$ ) est diluée dans  $50 \mu\text{L}$  de DMF (diméthylformamide) et ajoutée en excès à  $1 \text{ mL}$  de protéine concentrée à environ  $1 \text{ mg/mL}$ . La solution réactionnelle est incubée pendant  $2 \text{ h}$  à  $4^\circ\text{C}$ . Le tampon protéique contient du TCEP, un réducteur qui permet de réduire les ponts disulfures autorisant ainsi le groupement sulfhydryle de la cystéine à interagir avec le groupement maléimide de la sonde (Figure 26).

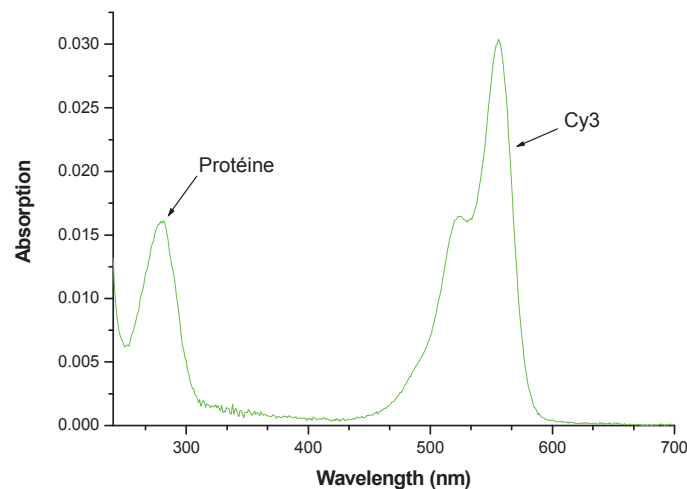


*Figure 26 : Réaction entre un composé maléimide et une molécule contenant un groupement sulfhydryl.*

Pour éliminer le fluorophore qui ne se serait pas lié à la protéine, on effectue une gel filtration (Superdex 200) sur une colonne équilibrée avec 25 mM Tris-HCl pH 7,5, 50 mM NaCl, 0,5 mM PMSF, 2 mM TCEP. Elle a été effectuée à l'IGBMC par Nicolas Lévy sur les 1 mL de protéine marquée.

Les fractions contenant la protéine d'intérêt sont identifiées par gel SDS-PAGE et la protéine est concentrée sur des filtres à centrifuger (Amicon15 MWCO 10 kDa). La protéine concentrée obtenue est aliquotée et conservée dans 5 % de glycérol à -80°C.

Pour calculer le pourcentage de protéine marquée, on observe son spectre d'absorption (Figure 27). L'absorption à 550 nm correspond au Cy3 ( $\epsilon = 150.000 \text{ M}^{-1}\text{cm}^{-1}$ ) et l'absorption à 280 nm correspond principalement à la protéine ( $\epsilon = 43.890 \text{ M}^{-1}\text{cm}^{-1}$ ). Le Cy3 ayant également une faible absorbance à 280 nm, un facteur correctif est appliqué pour déterminer l'absorbance réelle de la protéine :  $\text{Abs}_{280} \text{ réelle} = \text{Abs}_{280} \text{ observée} - (0,08 \times \text{Abs}_{\text{max}} \text{ du Cy3})$ . On peut ainsi calculer les concentrations respectives en Cy3 et en protéine et le ratio de leurs concentrations, la proportion de protéine marquée variant de 50 % à 70 %.



*Figure 27 : Spectre d'absorption de la protéine SRA (408-643) marquée au Cy3.*

## 2. Séquences oligonucléotidiques

### 2.1. Séquences utilisées

*Tableau 1 : Séquences des principaux oligonucléotides utilisés dans nos expériences.*

Oligonucléotides non-méthylés		Oligonucléotides héli-méthylés		Oligonucléotides bi-méthylés	
Noms	Séquences	Noms	Séquences	Noms	Séquences
NM	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'	HM	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'	BM	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'
NM3	5'-GGG <b>C</b> CC <b>C</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'	HM3	5'-GGG <b>C</b> CC <b>C</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'	BM3	5'-GGG <b>C</b> CC <b>C</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'
NM5	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'	HM5	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'	BM5	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'
NM6	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'				
NM7	5'-GGG CC <b>A</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'	HM7	5'-GGG CC <b>A</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'	BM7	5'-GGG CC <b>A</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'
NM8	5'-GGG CC <b>G</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'	HM8	5'-GGG CC <b>G</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'	BM8	5'-GGG CC <b>G</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'
NM9	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'	HM9	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'	BM9	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>G</b> GT CCC-5'
NM10	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>C</b> AT CCC-5'	HM10	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>C</b> AT CCC-5'	BM10	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>C</b> AT CCC-5'
NM11	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>A</b> GT CCC-5'	HM11	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>A</b> GT CCC-5'		
NM12	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>A</b> GT CCC-5'	HM12	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>A</b> GT CCC-5'	BM12	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>A</b> GT CCC-5'
NM13	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>A</b> GT CCC-5'	HM13	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>A</b> GT CCC-5'	BM13	5'-GGG CC <b>C</b> CA GGG-3' 3'-CCC GG <b>A</b> GT CCC-5'

Site de reconnaissance CpG ; M : cytosine méthylée ; A<sup>P</sup> : 2-aminopurine

La séquence de référence (gris sombre) comporte un site CpG central héli-méthylé (HM). Dans les deux autres séquences modèles (gris clair) ce site CpG est non-méthylé (NM) ou bi-méthylé (BM). Pour les expériences de FRET et d'anisotropie, ces oligonucléotides (HM, NM ou BM) sont marqués respectivement par du Cy5 ou de la fluorescéine à l'extrémité 5' du brin complémentaire. L'ensemble des autres séquences oligonucléotidiques est dérivé des oligonucléotides de référence (HM, NM ou BM) dans lesquels la 2-aminopurine (A<sup>P</sup>) remplace une des base naturelle à divers positions (3, 5, 6, 7, 8, 9, 5', 7', 8') au sein de la séquence.

L'oligonucléotide utilisé comme référence est un oligonucléotide hémi-méthylé (méthylé au niveau de la cytosine sur l'un des deux brins uniquement) de 12 paires de base. Sa séquence est présente dans le promoteur du gène *RBI* et est identique à celle utilisée par l'équipe de Dhe-Paganon pour l'élucidation de la structure tridimensionnelle du complexe SRA/ADN (Avvakumov et al., 2008). L'ensemble des autres séquences oligonucléotidiques est dérivé de cet oligonucléotide de référence et détaillé dans les Tableaux 1 et 2. Pour les expériences de titrage par transfert d'énergie non-radiatif (FRET), les oligonucléotides sont marqués à leur extrémité 5' par du Cy5. Pour les expériences d'anisotropie de fluorescence, ils sont marqués à leur extrémité 5' par de la fluorescéine-Ex, 5-isomère (FAM-EX-5). Enfin, les oligonucléotides utilisés pour l'étude détaillée de la dynamique moléculaire d'interaction de la protéine SRA avec l'ADN sont marqués avec de la 2-aminopurine, substituant un nucléotide de la séquence de l'oligonucléotide de référence au voisinage du site de liaison CpG de la protéine SRA. Tous ces oligonucléotides sont commandés chez IBA sous forme lyophilisée et hautement purifiée (double HPLC).

## 2.2. Mise en solution

Les séquences oligonucléotidiques délivrées sous forme lyophilisée sont reconstituées dans de l'eau millipore et conservées à -20°C. Les concentrations des solutions mères sont déterminées par lecture de l'absorbance à 260 nm, les coefficients d'extinction molaire des oligonucléotides étant compris entre 86.500 et 135.020 M<sup>-1</sup>cm<sup>-1</sup>.

## 2.3. Hybridation des oligonucléotides

Les oligonucléotides doubles brins sont préparés par hybridation au ratio 1/1 (mêmes quantités molaires pour les deux simples brins) afin d'éviter la présence d'ADN simple brin dans nos expériences. La concentration finale d'oligonucléotide double brin est généralement de 50 µM dans un tampon 25 mM Tris-HCl pH 7,5, 150 mM NaCl. Après ajout des quantités nécessaires d'oligonucléotides simples brins au tampon, le mélange est incubé pendant 2 min à 90°C. La solution est ensuite laissée à température ambiante toute la nuit à l'abri de la lumière pour permettre la diminution progressive de sa température et la fusion des deux brins

complémentaires. Une fois les oligonucléotides doubles brins formés, ils sont conservés à -20°C.

### 3. Conditions expérimentales des mesures des interactions protéine/ADN

L'ensemble des mesures d'interaction de la protéine SRA avec l'ADN est effectué dans du tampon 25 mM Tris-HCl pH 7,5, 50 mM NaCl, 0,5 mM PMSF, 2 mM TCEP, 0,04 % PEG 20.000. Le TCEP est choisi comme agent réducteur car, comparé aux autres réducteurs habituellement utilisés tel que le DTT (dithiothreitol) ou le  $\beta$ -mercaptoéthanol, il a l'avantage d'être moins odorant et plus résistant à l'oxydation par l'air, prévenant de ce fait l'apparition d'une bande d'absorption parasitant les mesures d'absorption de la protéine. Afin de prévenir l'adsorption de la protéine sur les parois des cuves de mesure en quartz, du polyéthylène glycol 20.000 (PEG) entre dans la composition du tampon de mesure. En recouvrant les parois des cuves, il permet de diminuer l'adsorption de la protéine en empêchant par conséquent la diminution artificielle de la concentration en protéine au centre de la cuve (Vuilleumier et al., 1997).

Toutes les expériences sont réalisées à 20°C à l'exception des courbes de fusion des oligonucléotides.

## B. Méthodes

### 1. Spectroscopie d'absorption UV-visible

#### 1.1. Principe et appareillage

Les spectres d'absorption sont enregistrés sur un spectrophotomètre UV-visible Cary 400 (Varian) à double faisceau avec une bande passante de 2 nm, dans des cellules thermostatées. Les cuves de mesure utilisées sont des cuves en quartz de parcours optique de



1 cm, transparentes dans l'UV et dans le visible. La ligne de base est enregistrée avec une cuve de référence et une cuve de mesure contenant chacune du tampon.

La spectrophotométrie d'absorption UV-visible consiste à mesurer l'absorption d'énergie par une molécule lors de son interaction avec des rayonnements UV (200 à 350 nm) ou visible (350 à 800 nm). A l'échelle macroscopique, l'absorbance  $A$  d'une solution de concentration  $c$  et d'épaisseur  $l$  est donnée par la loi de Beer-Lambert, le coefficient d'extinction  $\varepsilon$  de la molécule étudiée permettant de calculer sa concentration:

$$A = \varepsilon cl \quad (1)$$

Lorsque les solutions sont relativement concentrées, on observe une augmentation artéfactuelle de l'absorption due à la diffusion de la lumière. La contribution de la lumière diffusée est estimée selon la méthode de (Shih and Fasman, 1972) selon l'équation :

$$\log DO_d(\lambda) = A \log \lambda + B \quad (2)$$

où  $DO_d(\lambda)$  est l'absorbance due à la diffusion de la lumière à la longueur d'onde  $\lambda$ . Les coefficients  $A$  et  $B$  sont obtenus par ajustement de l'équation (2) aux points expérimentaux en utilisant la méthode des moindres carrés, à des longueurs d'onde où l'échantillon n'absorbe pas, et où l'absorbance mesurée résulte uniquement de la diffusion. Ainsi, le spectre d'absorption corrigé de la diffusion est obtenu par l'équation :

$$DO_{corr}(\lambda) = DO_{mes}(\lambda) - DO_d(\lambda) \quad (3)$$

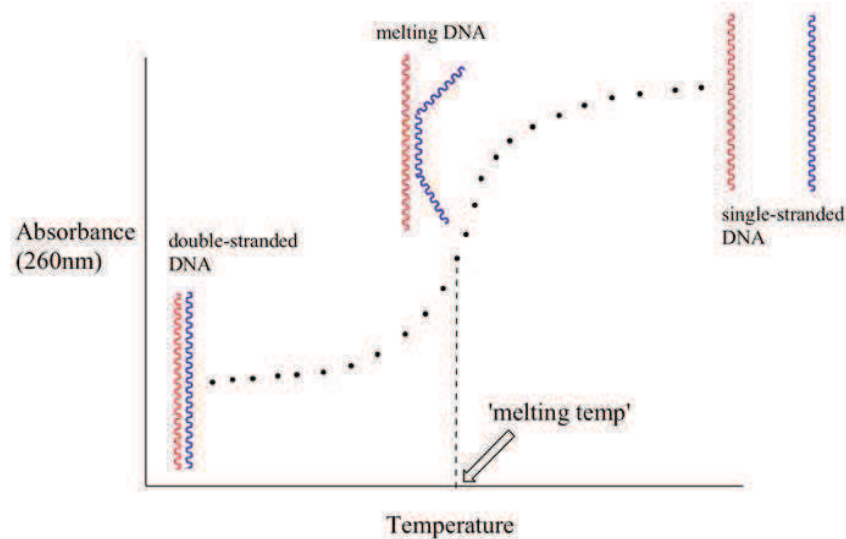
Les mesures d'absorbance de nos divers échantillons ont permis de calculer la concentration des oligonucléotides par lecture à 260 nm ( $\varepsilon$  des oligonucléotides simples brins compris entre 86.500 et 13.5020  $M^{-1}cm^{-1}$ ,  $\varepsilon$  des oligonucléotides doubles brins compris entre 178.641 et 188.543  $M^{-1}cm^{-1}$ ) ainsi que celle de la protéine par lecture à 280 nm ( $\varepsilon = 43.890 M^{-1}cm^{-1}$ ).

## 1.2. Courbes de fusion des oligonucléotides

Afin de comparer la stabilité des oligonucléotides doubles brins marqués à la 2-aminopurine à celle des oligonucléotides doubles brins de référence, les courbes de fusion des



divers oligonucléotides à une concentration de 2  $\mu\text{M}$  sont enregistrées par suivi de l'absorbance à 260 nm en fonction de la température (Figure 28).



*Figure 28 : Courbe de fusion théorique d'un ADN.*

L'absorbance des oligonucléotides est la somme de l'absorbance de leurs nucléotides additionnée de l'effet de leurs interactions. Dans l'ADN dans sa forme native en double hélice, l'empilement des bases est tel qu'il absorbe moins par nucléotide à 260 nm que dans sa forme dénaturée (chaînes individuelles). Ainsi, en suivant l'augmentation de l'absorbance des oligonucléotides à 260 nm en fonction de la température, on peut observer le passage de l'ADN double brin à l'ADN simple brin. L'augmentation de l'absorbance observée caractérise le phénomène d'hyperchromicité  $h_r$ , qui est défini par :

$$\%h_r = 100 \frac{A_{sb} - A_{db}}{A_{db}} \quad (4)$$

où  $A_{sb}$  est l'absorbance de l'ADN sous sa forme simple brin, et  $A_{db}$  est l'absorbance de l'ADN sous sa forme double brin (Bloomfield et al., 2000). Pour nos divers oligonucléotides, l'hyperchromicité observée est d'environ 10 %.

La courbe de fusion est également caractérisée par son point d'inflexion, qui correspond à la demi-variation d'absorbance. La température correspondante est la température de fusion, notée  $T_m$ , et caractérise la stabilité des oligonucléotides.

Il faut noter que le dépliement des formes simples brins de l'ADN peut également entraîner une variation de l'absorbance en fonction de la température. Ces variations doivent donc être prises en compte dans les mesures des courbes de fusion des oligonucléotides doubles brins lorsque l'un des simples brins composant le duplexe est présent en excès par rapport à son brin complémentaire. Une soustraction des absorbances correspondantes aux variations d'absorbance mesurées permet d'obtenir les variations réelles de l'ADN double brin.

## 2. Spectroscopie de fluorescence à l'état stationnaire

### 2.1. Emission de fluorescence

#### 2.1.1. Principe et appareillage

Les spectres d'émission de fluorescence sont enregistrés à l'aide de deux spectrofluorimètres, le FluoroMax-3 (Jobin Yvon) et le Fluorolog (Horiba, Jobin Yvon), tous deux équipés d'un porte-cuve thermostaté.

La source d'excitation des spectrofluorimètres est une lampe à xénon. La longueur d'onde d'excitation est sélectionnée par un monochromateur simple (FluoroMax-3) ou double (Fluorolog) réseau muni de deux fentes de largeur réglable. La fluorescence des solutions contenues dans des cuves en quartz est détectée à 90° à partir du centre de la cuve. Un monochromateur d'émission, également doté de deux fentes réglables, permet de sélectionner la longueur d'onde d'émission. La correction de l'instabilité du faisceau d'excitation due aux fluctuations de la lampe est réalisée à l'aide d'une photodiode calibrée. Les réponses du monochromateur et du photomultiplicateur à l'émission en fonction de la longueur d'onde sont corrigées par des facteurs déterminés par le constructeur.

En fluorescence, les mesures sont réalisées à angle droit. De ce fait, les mesures de fluorescence ne concernent que les molécules situées au centre de la cuve de mesure. Lorsque les solutions sont concentrées, l'absorbance est élevée et un grand nombre de photons sont absorbés dans les tous premiers millimètres de la cuve. Il en résulte que l'intensité incidente

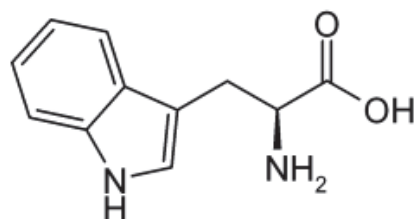
parvenant au centre de la cuve est faible et donc n'excitera qu'un nombre restreint de fluorophores. La fluorescence mesurée est, de ce fait, sous évaluée. Ce phénomène correspond à l'effet d'écran. La réduction artéfactuelle de l'intensité de fluorescence peut être corrigée en multipliant l'intensité de fluorescence mesurée à chaque longueur d'onde d'excitation  $\lambda_{ex}$  par le facteur  $CF$  (Nordlund et al., 1993; Xu and Nordlund, 2000):

$$CF = \frac{2.303 \times A(\lambda_{ex})}{1 - 10^{-A(\lambda_{ex})}} \quad (5)$$

La protéine SRA (408-648) ayant tendance à s'adsorber sur les parois des cuves de mesure, un certain nombre de précautions ont été prises lors de la préparation des divers échantillons. Ainsi, outre l'addition de PEG 20.000 au tampon utilisé pour les mesures (voir p. 86), toutes les solutions mesurées sont préparées extemporanément dans des tubes Eppendorf possédant une faible adhésion de surface ("low binding"). En effet, l'adsorption constitue une limitation majeure pour les études de la protéine à des concentrations inférieures au micromolaire puisqu'elle entraîne une diminution artéfactuelle de la concentration en protéine au centre de la cuve et par conséquent de l'intensité de fluorescence détectée.

### 2.1.2. Fluorescence du tryptophane

Au sein des protéines, les trois acides aminés aromatiques, phénylalanine, tyrosine et tryptophane (Trp), sont tous fluorescents. Parmi ces acides aminés, le plus intéressant est le Trp car il possède des propriétés d'absorption et d'émission supérieures à celles des deux autres acides aminés aromatiques. De plus, le Trp peut être excité sélectivement et il est très sensible aux propriétés physico-chimiques de son environnement.



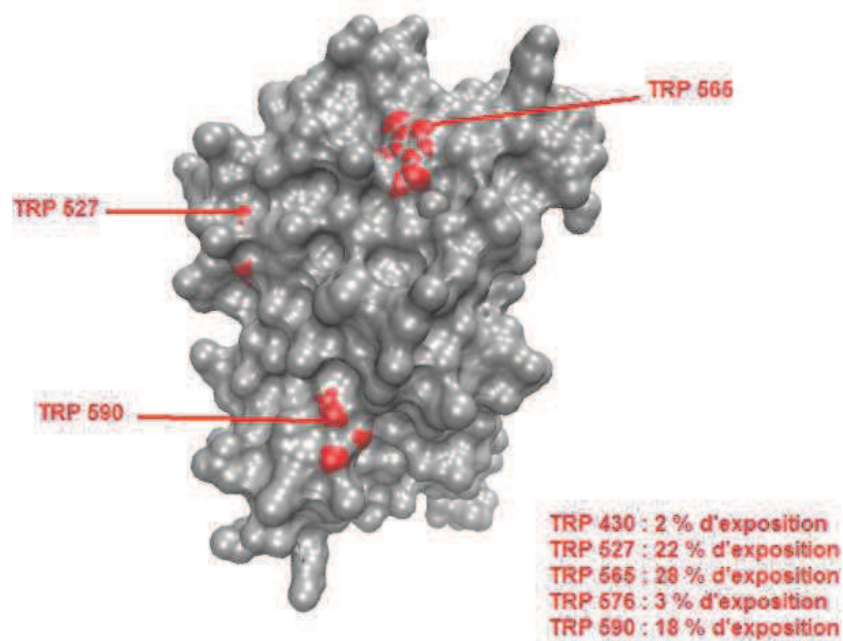
*Figure 29: Structure chimique du tryptophane.*

La fluorescence du Trp est due au noyau indole constituant sa chaîne latérale (Figure 29). Son maximum d'émission, situé à 350 nm dans l'eau à pH neutre, est fortement dépendant de la polarité du milieu et de l'environnement local. Ainsi, pour un Trp dans un environnement complètement apolaire, un déplacement du maximum d'émission vers le bleu est observé. Le Trp possède un rendement quantique de 0,14 dans l'eau (Eisinger and Navon, 1969). A pH neutre, le déclin de fluorescence du Trp est caractérisé par deux durées de vie, une composante dominante de 3 ns et une plus courte de 0,6 ns. Cette hétérogénéité est due à l'existence d'isomères conformationnels rotationnels du Trp à l'état fondamental, nommés rotamères.

L'émission de fluorescence des protéines est gouvernée par la fluorescence de ses résidus Trp, qui peuvent être sélectivement excités à 295-300 nm et qui, selon leur localisation, contribuent de manière inégale à l'émission totale de fluorescence de la protéine. Les résidus Trp étant sensibles à l'environnement, l'interprétation de leur spectre d'émission peut donner des informations sur l'emplacement des résidus au sein de la protéine. Ainsi, l'émission d'un résidu Trp exposé en surface apparaîtra à de plus grandes longueurs d'onde que celle d'un résidu Trp enfoui dans la protéine. Par conséquent, l'émission et le rendement quantique du Trp peuvent grandement varier d'une protéine à l'autre, donnant des informations sur la structure des protéines, et permettant de suivre leur repliement. Bien que la variabilité du rendement quantique et des temps de vie des protéines soit due à un certain nombre d'interactions qui dépendent toutes de la structure de la protéine, il n'a pas été observé de corrélation claire entre les valeurs des temps de vie de fluorescence et du rendement quantique, et la structure des protéines. Dans la plupart des cas, le déclin de fluorescence du Trp dans les protéines est multi-exponentiel. De même qu'en solution, l'hétérogénéité des déclins de fluorescence du Trp dans les protéines est due en général à un équilibre conformationnel à l'état fondamental (modèle des rotamères).

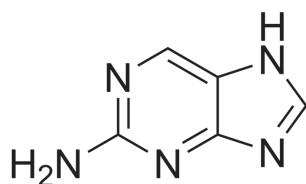
La protéine SRA (408-643) possède cinq résidus Trp (Figure 22) dont trois sont relativement exposés au solvant (Figure 30). Les caractéristiques spectroscopiques des résidus Trp de la protéine tels que leurs spectres d'absorption, d'excitation et d'émission, ainsi que leurs rendements quantiques et leurs temps de vie de fluorescence, permettent de caractériser l'état conformationnel de la protéine et de confirmer son bon repliement après production et purification. Toutes ces mesures sont effectuées avec une concentration en protéine de 0,4  $\mu\text{M}$ , excepté pour la mesure du rendement quantique où la concentration est de 2  $\mu\text{M}$ . Les

spectres d'émission de la protéine sont enregistrés avec une longueur d'onde d'excitation de 295 nm pour exciter sélectivement les résidus Trp.



*Figure 30: Degrés d'exposition des résidus Trp de la protéine SRA (408-643). Seul trois résidus Trp sont visibles dans la représentation de surface du domaine SRA. Les deux autres Trp sont enfoui dans la protéine. Calculs effectués sur "<http://curie.utmb.edu/getarea.html>" et représentation sur VMD (code PDB 2PB7).*

### 2.1.3. Fluorescence de la 2-aminopurine



*Figure 31: Structure chimique de la 2-aminopurine.*

La 2-aminopurine est un analogue fluorescent de l'adénine (Figure 31). En solution, la 2-Ap possède un maximum d'absorption à 303 nm et un maximum d'émission autour de 370 nm, avec un rendement quantique élevé de 0,68 et un temps de déclin de fluorescence monoexponentiel proche de 10 ns. La 2-Ap forme des interactions Watson-Crick stables avec

la thymine, et des paires de bases mésappariées avec la cytosine, l'adénine et la guanine. Elle peut donc être incorporée dans un ADN double brin à la place d'une base naturelle et y jouer le rôle de rapporteur fluorescent. En effet, la 2-Ap est sensible à son microenvironnement et plus particulièrement à l'empilement (stacking) avec ses bases voisines. Pour cette raison elle donne des informations locales au sein d'un acide nucléique. Dans l'ADN, la fluorescence de la 2-Ap est partiellement éteinte et son déclin devient complexe avec plusieurs temps de vie de fluorescence qui correspondent à des états conformationnels différents, traduisant l'hétérogénéité de l'environnement de la 2-Ap. De même, le déclin d'anisotropie résolu en temps présente deux à trois temps de corrélation de rotation qui révèlent la flexibilité de l'oligonucléotide.

Dans nos expériences, l'introduction de la 2-Ap à différentes positions dans nos oligonucléotides de référence nous a permis d'élucider les changements de conformation locaux induits par la liaison de la protéine SRA (408-643). La variation du rendement quantique indique un éventuel changement de conformation au voisinage de la 2-Ap. De manière plus précise, les mesures de temps de vie de fluorescence de la 2-Ap donnent des informations sur les différents états conformationnels de la 2-Ap, et l'anisotropie résolue en temps sur les variations de la flexibilité de l'ADN en présence de protéine. Les mesures sont effectuées sur 3  $\mu\text{M}$  d'oligonucléotides en absence ou en présence de 4  $\mu\text{M}$  de protéine. Une longueur d'onde d'excitation de 315 nm permet d'exciter sélectivement la 2-Ap.

#### 2.1.4. Mesure du rendement quantique de fluorescence

Le rendement quantique d'une molécule correspond au rapport du nombre de photons émis au nombre de photons absorbés et permet de caractériser un fluorophore dans son environnement. Pour estimer le rendement quantique d'un fluorophore, on le compare à une molécule de référence dont le rendement quantique est connu. En ce qui concerne la mesure du rendement quantique des résidus Trp, la référence communément choisie est le N-acétyl-L-tryptophanamide (NATA) ( $\Phi_{\text{ref}} = 0.14$ ). Pour la 2-aminopurine, la référence est la 2-aminopurine riboside ( $\Phi_{\text{ref}} = 0.68$ ) qui est soluble dans les solvants aqueux.

Le rendement quantique d'un composé est calculé en utilisant la formule suivante :

$$\Phi = \frac{I \times Abs_{ref}}{I_{ref} \times Abs} \times \Phi_{ref} \quad (6)$$

où  $\Phi$  est le rendement quantique,  $I$  l'aire sous la courbe du spectre d'intensité de fluorescence du composé et  $Abs$  la valeur d'absorption à la longueur d'onde d'excitation de ce même composé. L'indice *ref* se rapporte au fluorophore de référence de rendement quantique connu. Si l'absorption des échantillons est supérieure à 0,05, l'intensité de fluorescence mesurée est corrigée par l'effet d'écran.

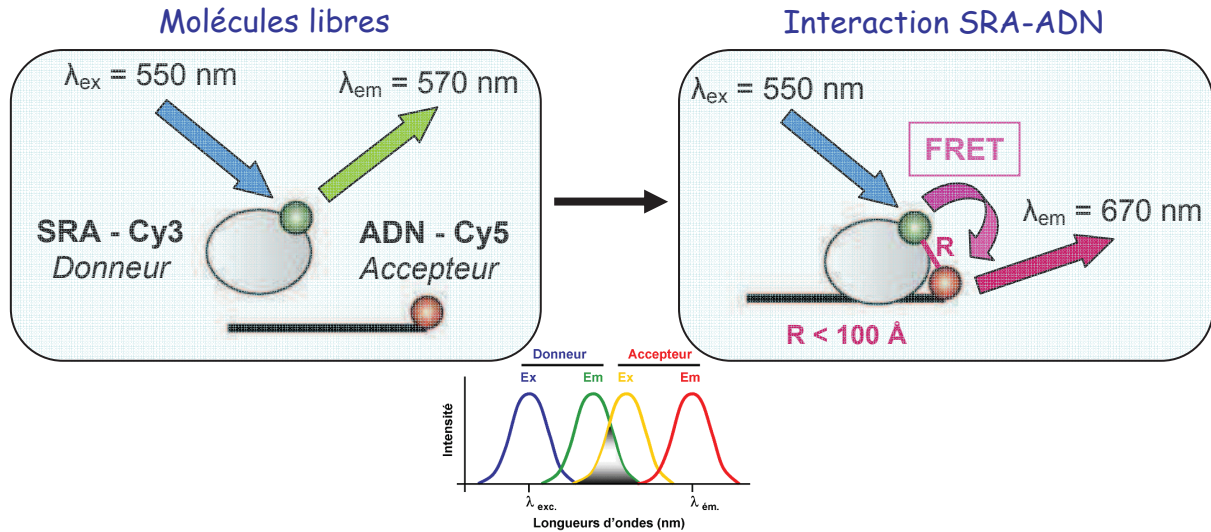
Dans le cas du calcul du rendement quantique de la 2-aminopurine des oligonucléotides double brins marqués, l'absorbance de la 2-Ap à 303 nm est polluée par l'absorbance résiduelle des oligonucléotides à cette longueur d'onde. La valeur d'absorbance imputable à la 2-Ap n'est donc pas directement lisible sur le spectre d'absorption. Dans ce cas, l'échantillon et la référence sont mesurés à la même concentration et donc à la même valeur d'absorbance. Alors, la formule initiale de calcul du rendement quantique devient :

$$\Phi = \frac{I}{I_{ref}} \times \Phi_{ref} \quad (7)$$

### 2.1.5. Mesure de l'interaction de la protéine SRA aux acides nucléiques par transfert d'énergie non-radiatif (FRET)

Le transfert d'énergie est un phénomène observé lorsqu'une molécule luminescente (donneur) cède, à l'état excité, son énergie à une seconde molécule (accepteur). Cette dernière peut alors se désexciter en émettant un photon de fluorescence. Le FRET (Fluorescence Resonance Energy Transfer) correspond à un transfert d'énergie non radiatif, c'est-à-dire sans émission de photons, par couplage dipôle-dipôle. Il est possible si le spectre d'émission du donneur recouvre en partie celui de l'accepteur (Figure 32). L'efficacité du transfert d'énergie dépend de l'importance du recouvrement du spectre d'émission du donneur avec le spectre d'absorption de l'accepteur, du rendement quantique du donneur, de l'orientation relative des dipôles du donneur et de l'accepteur, et de la distance entre le donneur et l'accepteur (Lakowicz, 1999).





*Figure 32: Principe du FRET : exemple de l'interaction de la protéine SRA (408-643) marquée au Cy3 avec de l'ADN marqué au Cy5.*

Une des raisons pour lesquelles cette méthode est si répandue dans les études biophysiques est que le FRET sonde efficacement des distances comprises entre 10 et 100 Å environ (soit de 1 à 10 nm), ce qui correspond à l'ordre de grandeur de la taille d'une protéine ou de l'épaisseur d'une membrane cellulaire. Une des applications les plus communes du FRET est la détermination de la distance entre deux molécules individuellement marquées par un donneur et un accepteur.

La distance entre le donneur et l'accepteur peut être déterminée à partir de l'efficacité de transfert  $E$ , qui correspond à la fraction de photons absorbés par le donneur qui sont transférés à l'accepteur. Cette fraction est déterminée par :

$$E = \frac{k_T(r)}{\tau_D^{-1} + k_T(r)} \quad (8)$$

On rappelle que la vitesse du transfert d'énergie  $k_T(r)$  s'exprime selon la formulation de Förster (Förster, 1948) :

$$k_T(r) = \frac{1}{\tau_D} \left( \frac{R_0}{r} \right)^6 \quad (9)$$

ce qui implique que :

$$E = \frac{R_0^6}{R_0^6 + r^6} \quad (10)$$

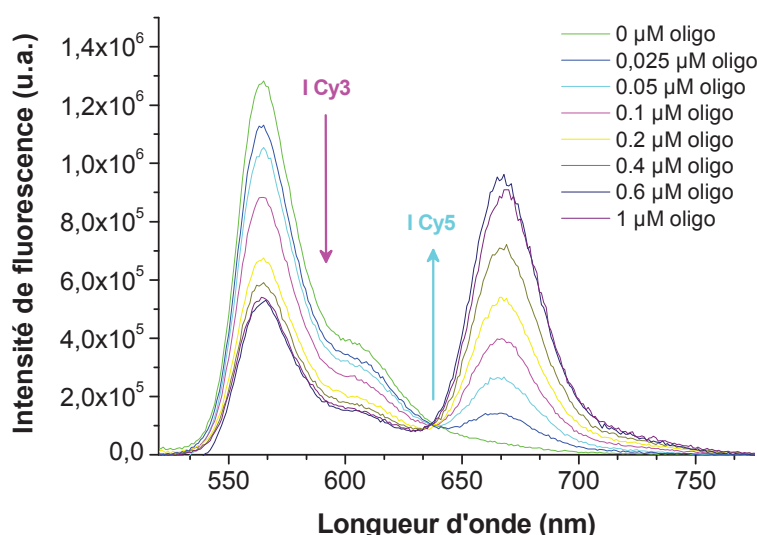
où  $\tau_D$  est le temps de vie du donneur en absence de l'accepteur,  $R_0$  la distance de Förster, et  $r$  la distance entre le donneur et l'accepteur.  $R_0$  est défini comme la distance pour laquelle on a 50 % d'efficacité de transfert d'énergie. Sa valeur peut être prédite à partir des propriétés spectrales du donneur et de l'accepteur et est caractéristique d'un couple de fluorophores donné. Pour le couple de chromophores Cy3/Cy5,  $R_0$  est égal à 53 Å (Stein et al., 2011).

En pratique, lorsqu'il n'y a pas de décalage du spectre d'émission du donneur suite à la liaison de l'accepteur,  $E$  est mesurée à partir de l'intensité de fluorescence relative du donneur en absence ( $F_D$ ) et en présence d'accepteur ( $F_{DA}$ ) :

$$E = 1 - \frac{F_{DA}}{F_D} \quad (11)$$

Ainsi, la distance interchromophore  $r$  est calculée à partir des valeurs de  $R_0$  et  $E$  grâce à l'équation :

$$r = R_0 \left( \frac{1}{E} - 1 \right)^{1/6} \quad (12)$$



*Figure 33: Mesure de l'interaction protéine SRA/ADN par FRET. Une concentration croissante d'oligonucléotides (oligo) marqués au Cy5 est ajoutée à une concentration fixe de 0,05 μM de protéine SRA (408-643) marquée au Cy3. On observe une diminution progressive de l'émission du donneur Cy3 parallèlement à l'augmentation de l'émission de l'accepteur Cy5, témoignant de la liaison des oligonucléotides au domaine SRA. La longueur d'onde d'excitation est de 500 nm. L'émission du Cy5 est corrigée par l'excitation directe du fluorophore à cette longueur d'onde d'excitation.*

Le FRET est donc un outil intéressant pour sonder des distances au sein de biomolécules doublement marquées. D'autre part, l'association de molécules peut également être observée par FRET, ce qui permet de suivre une interaction, même sans calcul de distance. Dans le cadre des travaux effectués dans cette thèse, nous avons utilisé la technique de FRET pour suivre l'interaction de la protéine SRA (408-643) avec différents types d'ADN. La protéine SRA (408-643) est marquée sélectivement par du Cy3, et les oligonucléotides par du Cy5 (voir p. 80 et 84). Le couple Cy3/Cy5 présente une valeur  $R_0$  de 53 Å (Stein et al., 2011) qui permet de suivre l'interaction protéine/ligand et de calculer avec précision les distances interchromophores. Dans nos expériences, la technique de FRET est utilisée pour suivre la liaison d'oligonucléotides marqués au Cy5 ajoutés en concentration croissante (de 0 à 2 µM) à une concentration fixe de 0,05 µM de protéine SRA (408-643) marquée au Cy3. Lors du titrage, le phénomène de FRET est mis en évidence par la diminution progressive de l'émission du donneur Cy3 parallèlement à l'augmentation de l'émission de l'accepteur Cy5 (Figure 33). L'efficacité de transfert, calculé par  $E = 1 - F_{DA}/F_D$  (équation 11), est tracée en fonction de la concentration en oligonucléotides. La courbe théorique est ajustée aux points expérimentaux par l'équation adaptée du modèle de (Didier et al., 2011) pour une liaison protéine/ligand de stoechiométrie 1/1 :

$$E = E_f \left\{ \frac{1 + K_a ([SRA]_{tot} + [ADN]_{tot}) - \left\{ [1 + K_a ([SRA]_{tot} + [ADN]_{tot})]^2 - 4K_a^2 [SRA]_{tot} [ADN]_{tot} \right\}^{1/2}}{2K_a [SRA]_{tot}} \right\} \quad (13)$$

où  $[SRA]_{tot}$  et  $[ADN]_{tot}$  sont les concentrations totales en protéine SRA (408-643) et en oligonucléotide, respectivement. A saturation, la valeur finale de l'efficacité de FRET,  $E_f$ , donne la distance interchromophore moyenne de tous les complexes en solution. Les points expérimentaux ont été ajustés à ce modèle afin de déterminer la constante d'affinité  $K_a$  de la protéine SRA (408-643) pour les différents types d'ADN utilisés. Les mesures de FRET sont effectuées à une longueur d'onde d'excitation de 500 nm pour exciter sélectivement le Cy3, et enregistrées de 520 à 800 nm pour observer les spectres du donneur Cy3 et de l'accepteur Cy5. L'émission du Cy5 est corrigée de la contribution de fluorescence de ce résidu, suite à son excitation directe à 500 nm. Cette correction est réalisée par des mesures contrôles de spectres d'émission de fluorescence des oligonucléotides marqués au Cy5 ajoutés en concentration croissante à la protéine SRA (408-643) non marquée. Pour éviter l'adsorption de la protéine sur les parois des cuves de mesure, chaque point de la courbe de titrage est

préparé séparément dans un tube Eppendorf possédant une faible adhésion de surface ("low binding") (voir p. 90).

## 2.2. Anisotropie de fluorescence

### 2.2.1. Principe et appareillage

Une molécule fluorescente est caractérisée dans son état fondamental par un moment dipolaire de transition d'absorption  $\mu_A$ . Lorsque l'on illumine une population de fluorophores en solution par une lumière linéairement polarisée (c'est-à-dire caractérisée par un champ électrique qui vibre dans une direction unique), ceux dont les moments dipolaires de transition d'absorption  $\mu_A$  sont orientés dans une direction proche de celle du vecteur champ électrique du faisceau incident sont préférentiellement excités. C'est le phénomène de photosélection. La probabilité d'absorption de chaque molécule est proportionnelle au cosinus carré de l'angle  $\theta$  entre son vecteur de moment d'absorption  $\mu_A$  et le vecteur champ électrique incident.

La lumière de fluorescence émise est orientée selon un vecteur champ électrique parallèle à la direction du moment dipolaire de transition d'émission,  $\mu_E$ , qui caractérise la molécule fluorescente dans son état excité. Comme la distribution des molécules à l'état excité est anisotrope, du fait de la photosélection, il en résulte une lumière de fluorescence polarisée, en fonction de l'orientation moyenne de la population de molécules excitées au moment de l'émission.

Ce phénomène peut être quantifié par l'anisotropie de fluorescence,  $r$ . Les mesures d'anisotropie donnent des informations notamment sur la mobilité, la taille, la forme et la flexibilité des molécules. Plus le mouvement d'une molécule est rapide, plus son anisotropie est faible. Dans le cas d'une molécule immobilisée, on parle d'anisotropie fondamentale,  $r_0$ , définie par :

$$r_0 = (3 \cos^2 \alpha - 1) / 5 \quad (14)$$

où  $\alpha$  est l'angle entre les moments de transition d'absorption  $\mu_A$  et d'émission  $\mu_E$ .

Cette valeur de  $r_0$  est la valeur maximale pour l'anisotropie de la molécule considérée. Dans le cas particulier où les moments de transitions d'absorption et d'émission sont parallèles,  $r_0$  atteint une valeur limite égale à 0,4.

Lorsque la molécule fluorescente est en mouvement par suite de mouvements Browniens de rotation, il en résulte un changement de direction de son moment d'émission  $\mu_E$ , d'un angle  $\omega(t)$ , ceci pendant la durée de vie de l'état excité. Le mouvement provoquera une diminution de l'anisotropie de distribution des molécules fluorescentes résultant de la photosélection, et par conséquent induira une dépolarisation partielle ou totale de la fluorescence, en fonction de l'amplitude du mouvement. La dépolarisation peut aussi être induite par du transfert d'énergie d'une molécule « donneur » à une molécule « accepteur » orientée différemment.

Dans le cas général de mouvements Browniens,  $r$  est égal à :

$$r = (3 \cos^2(\alpha + \omega(t)) - 1) / 5 \quad (15)$$

où  $\alpha$  est l'angle entre les moments d'absorption  $\mu_A$  et d'émission  $\mu_E$ , et  $\omega(t)$  est l'angle de rotation Brownienne entre l'orientation du vecteur d'émission  $\mu_E$  au temps zéro d'excitation et son orientation au moment de l'émission du photon de fluorescence, c'est-à-dire à terme du temps passé dans l'état excité.

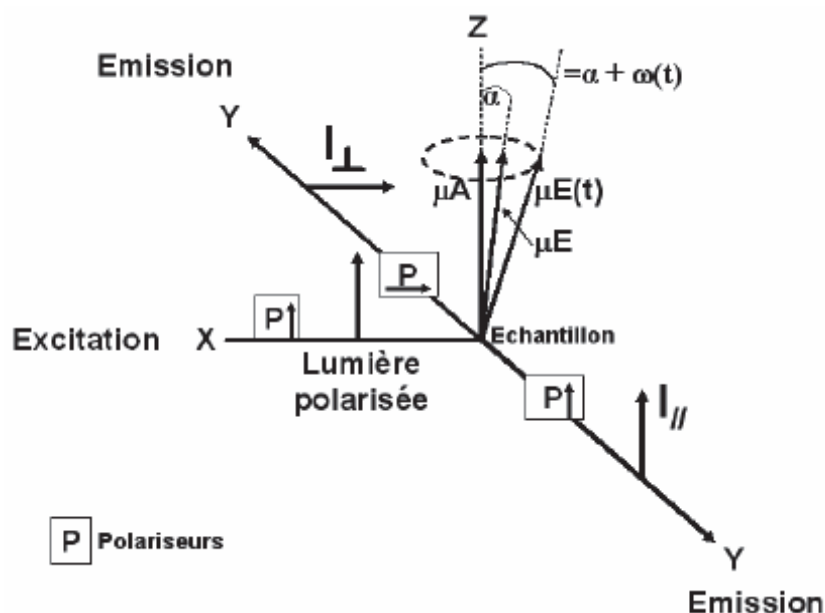


Figure 34: Schéma de la configuration en T pour la mesure d'anisotropie de fluorescence.

Pour les mesures d'anisotropie de fluorescence, la lumière incidente est habituellement orientée verticalement (selon l'axe Z) (Figure 34). Dans le cas d'un montage en T (qui s'oppose à un montage en L), l'émission de fluorescence est analysée à l'aide de deux polariseurs placés sur l'axe Y, l'un étant orienté parallèlement et l'autre perpendiculairement par rapport à l'axe Z. Ainsi on a  $I_Z = I_{//}$  et  $I_X = I_Y = I_{\perp}$ . L'intensité de fluorescence totale est égale à  $I = I_X + I_Y + I_Z = I_{//} + 2 I_{\perp}$ .

L'anisotropie de fluorescence,  $r$ , est alors déterminée par :

$$r = \frac{I_{//} - I_{\perp}}{I_{//} + 2I_{\perp}} \quad (16)$$

où  $I_{//}$  et  $I_{\perp}$  désignent les composantes de fluorescence respectivement parallèle et perpendiculaire à la direction de vibration de la lumière d'excitation verticale.

Les mesures d'anisotropie à l'état stationnaire sont réalisées à 20°C à l'aide d'un spectrofluorimètre SLM 8000 en format T équipé d'un porte-cuve thermostaté.

### 2.2.2. Mesure de la liaison des protéines SRA aux acides nucléiques par titrage direct

La liaison de la protéine SRA aux acides nucléiques peut être suivie par anisotropie de fluorescence en utilisant des oligonucléotides marqués à la fluorescéine. Dans le cas d'un titrage direct, une concentration fixe d'oligonucléotides, ici 0,2  $\mu\text{M}$ , est titrée par une concentration croissante de protéine. L'anisotropie de l'oligonucléotide en fonction de la concentration en protéine est ajustée par l'équation (Didier et al., 2011) :

$$r = r_0 + \frac{(r_f - r_0)}{n} \left( \frac{(1 + K_a(n[ADN]_{tot} + [SRA]_{tot})) - \sqrt{(1 + K_a(n[ADN]_{tot} + [SRA]_{tot}))^2 - 4nK_a^2[ADN]_{tot}[SRA]_{tot}}}{2K_a[ADN]_{tot}} \right) \quad (17)$$

où  $[SRA]_{tot}$  et  $[ADN]_{tot}$  désignent les concentrations totales en protéine SRA (408-643) et en oligonucléotide, respectivement.  $r_f$  représente l'anisotropie au plateau, lorsque tous les oligonucléotides sont saturés en protéine, tandis que  $r_0$  et  $r$  correspondent aux valeurs d'anisotropie de l'oligonucléotide marqué, respectivement en absence et en présence de protéine. Les points expérimentaux sont ajustés afin de déterminer la constante d'affinité  $K_a$

de la protéine SRA (408-643) pour les différents types d'ADN utilisés.  $n$  correspond au nombre de site de liaison. Comme l'oligonucléotide utilisé comme référence contient uniquement un site hémi-méthylé qui correspond au site de reconnaissance de la protéine SRA, on suppose  $n$  égal à 1. L'anisotropie des oligonucléotides marqués à la fluorescéine est enregistrée à une longueur d'onde d'émission supérieure à 500 nm à grâce à des filtres passe-haut (Kodak). La longueur d'onde d'excitation est de 470 nm. Afin de déceler un éventuel phénomène de diffusion qui pourrait être le témoin d'une agrégation des complexes protéine/ADN, les spectres d'absorption des solutions sont enregistrés entre 240 et 800 nm. D'autre part, leur intensité de fluorescence est également mesurée. Comme pour les mesures de FRET, pour éviter l'adsorption de la protéine sur les parois des cuves de mesure, chaque point de la courbe de titrage est préparé séparément dans un tube Eppendorf possédant une faible adhésion de surface ("low binding") (voir p.90).

### 3. Spectroscopie de fluorescence résolue en temps

La durée de vie de fluorescence d'une molécule est définie comme le temps moyen pendant lequel le fluorophore reste à l'état excité avant de revenir à son état fondamental. Cette durée de vie peut être modifiée lors de l'interaction de la molécule avec son environnement proche.

Les principales techniques de mesure de déclin de fluorescence d'un échantillon sont la technique de phase et de modulation (Gratton and Limkeman, 1983; Lakowicz and Maliwal, 1985), et la technique impulsionnelle selon la méthode du comptage de photon unique (Badea and Brand, 1979). La première technique repose sur une illumination continue de l'échantillon à l'aide d'un faisceau dont l'intensité est modulée sinusoïdalement en fonction du temps. La seconde méthode repose sur une excitation pulsée et une détection directe du temps de séjour de la molécule à l'état excité. Ces deux méthodes possèdent des résolutions dans des échelles de temps de la nanoseconde à la picoseconde, selon les sources et les détecteurs employés. Les mesures que nous avons réalisées dans le cadre de ces travaux ont été faites par la technique du photoélectron unique (PEU).



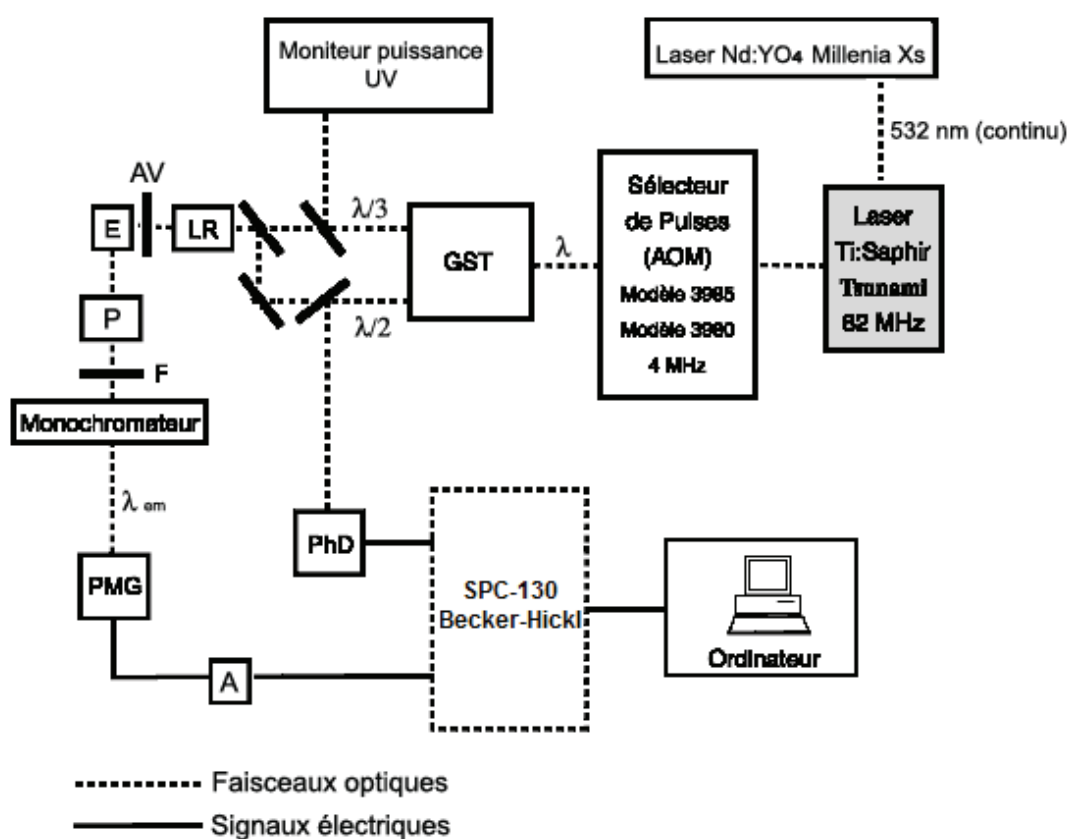
### 3.1. Principes de la technique du photoélectron unique et dispositif de mesure

La technique du PEU consiste à exciter des fluorophores par un laser pulsé dont l'intensité doit être suffisamment faible pour qu'au maximum une molécule soit excitée à chaque pulse. On considère que cette condition est remplie si l'on détecte moins de 5 photons de fluorescence pour 100 impulsions d'excitation. L'intervalle de temps entre l'impulsion initiale et la détection du photon, qui correspond au temps passé par une molécule à l'état excité, est mesuré. En répétant un grand nombre de fois cette mesure, il est possible de reconstituer un déclin de fluorescence.

La source d'excitation est constituée par un laser à saphir dopé au titane  $\text{Al}_2\text{O}_3:\text{Ti}^{3+}$  (Tsunami Spectra Physics) à modes bloqués (Figure 35). Ce laser est pompé par un laser continu solide néodyme yttrium vanadate ( $\text{Nd}:\text{YVO}_4$ ) modèle Millennia X (Spectra-Physics) émettant à 532 nm avec une puissance de sortie de 7 à 9 W. Le laser Ti:Saphir émet dans un intervalle de longueurs d'onde allant de 720 à 1080 nm, en délivrant des impulsions de largeur temporelle de 1 ps. La fréquence de répétition naturelle du laser est de 82 MHz. Cette fréquence, trop élevée pour la plupart des mesures de temps de vie de fluorescence des fluorophores usuels, est ramenée à 4 MHz à l'aide d'un sélecteur d'impulsions (pulse-picker). A la sortie de ce sélecteur, le faisceau est dirigé sur un système de cristaux non linéaires en accord de phase angulaire (GST). Dans un premier temps, le faisceau est focalisé sur un cristal LBO (lithium borate) permettant la formation de l'harmonique secondaire ( $\lambda/2$ ). Par la suite, après passage par une ligne de retard optique, le faisceau fondamental et le faisceau doublé sont focalisés sur un cristal BBO ( $\beta$  barium borate) permettant la formation de la troisième harmonique ( $\lambda/3$ ). Le doublage et le triplage de fréquence du faisceau fondamental permettent d'exciter un grand nombre de fluorophores usuels, respectivement dans la gamme de 420 à 480 nm pour le doublé (fluorescéine, rhodamine, ...) et dans la gamme de 280 à 320 nm pour le triplé (tryptophane, 2-aminopurine, ...).

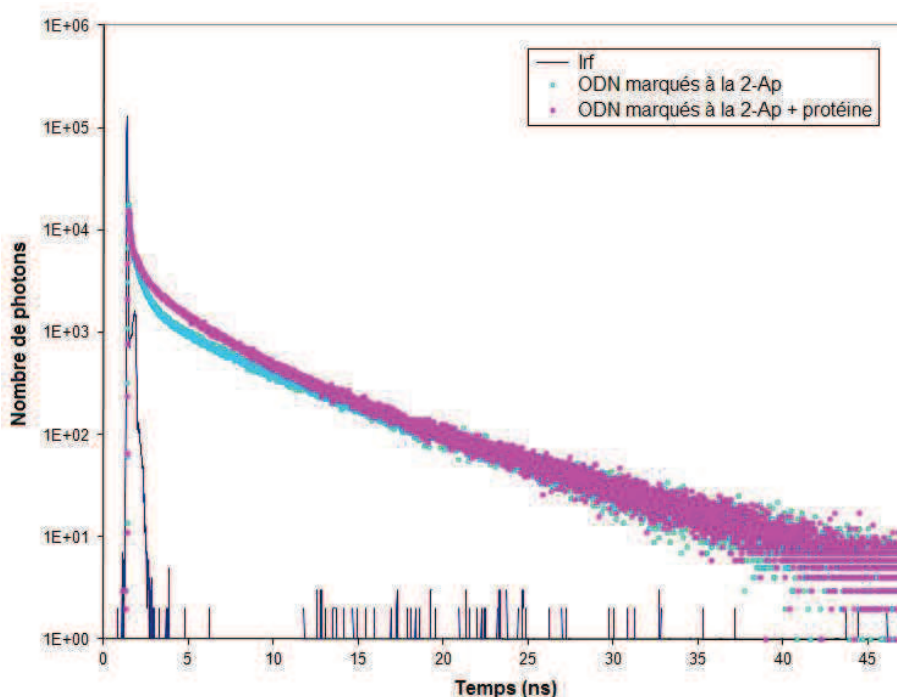
Une partie du faisceau doublé est envoyé vers une photodiode rapide (Hamamatsu) connectée à une voie du discriminateur rapide (Tennelec TC 454), générant le signal de départ (Start) pour le convertisseur temps-amplitude (TAC) (SPC-130 Becker-Hickl) et le début de la charge du condensateur. Les photons de fluorescence émis par l'échantillon sont détectés selon une direction perpendiculaire à l'excitation par un photomultiplicateur (PMT) à galette

de microcanaux (Hamamatsu R3809U) couplé à un préamplificateur (HFAC Becker-Hickl) (A). Les longueurs d'ondes d'émission sont sélectionnées par un monochromateur (Jobin-Yvon) dont la bande passante est ajustable (4 à 16 nm). Un polariseur (P) réglé à l'"angle magique" de  $54,7^\circ$  permet de rendre le signal de fluorescence indépendant des mouvements de rotation des molécules excitées. Le signal correspondant au premier photon détecté sert de signal "Stop" pour le TAC, ce qui provoque l'arrêt du chargement du condensateur. La charge accumulée dans le TAC est donc proportionnelle à l'intervalle de temps entre l'impulsion d'excitation (Start) et d'émission du photon (Stop). Cette charge est convertie en une tension comprise entre 0 et 10 V, qui est analysée dans un analyseur multicanaux (SPC-130 Becker-Hickl). 4096 canaux sont utilisés pour l'analyse de l'amplitude, ce qui, combiné à une échelle de temps de 50 ns du TAC, donne une linéarité de 12,20 ps/canal. L'analyseur multicanaux est piloté par un ordinateur à l'aide d'une interface électronique gérée par le logiciel SPCM.



*Figure 35 : Schéma du dispositif de mesure de fluorescence résolue en temps selon la méthode du comptage du photoélectron unique. GST : Générateur de seconde et triple harmonique, LR : lame rotatoire, AV : Atténuateur variable, P : Polariseur, PhD : Photodiode rapide, E : Echantillon, F : Filtre, PMG : Photomultiplicateur à galettes de microcanaux, A : Préamplificateur d'impulsions, SPC-130 : Module de comptage de photons uniques corrélés en temps.*

La mesure d'un seul photon n'est pas suffisante pour décrire le déroulement temporel de la fluorescence. L'historgramme de la probabilité d'émission de fluorescence en fonction du temps (Figure 36) est donc reconstitué par la mesure d'un grand nombre de photons, typiquement  $1 \times 10^6$  photons.



*Figure 36 : Déclin de temps de vie de fluorescence d'un oligonucléotide marqué à la 2-Ap ( $3 \mu\text{M}$ ) en absence ou en présence de protéine SRA (408-643) ( $4 \mu\text{M}$ ). La fonction de réponse (Irf) de l'appareil est également représentée. Les longueurs d'onde d'excitation et d'émission sont de 315 et 370 nm respectivement.*

Il est nécessaire d'enregistrer également la fonction de réponse qui permet de prendre en compte les imperfections de l'instrument (largeur temporelle d'excitation, incertitude de déclenchement de l'électronique, réponse du détecteur). Pour cela, on remplace l'échantillon par un réflecteur en aluminium poli (qui n'a pas de durée de vie de fluorescence), on amène la largeur d'onde d'observation à celle de l'excitation, et on veille à atténuer très fortement l'excitation. La largeur à mi-hauteur de cette réponse instrumentale est de l'ordre de 50 ps.

### 3.2. Mesure des durées de vie de fluorescence

Les mesures de temps de vie de fluorescence sont faites à l'"angle magique" de  $54,7^\circ$  pour éviter la contribution des mouvements de rotation des molécules excitées. Le déclin de fluorescence mesuré,  $I(t)$ , est le produit de convolution de la loi théorique d'émission de fluorescence de l'échantillon  $i(t)$  et de la fonction de réponse instrumentale effective  $E(t)$  :

$$I(t) = E(t) \times i(t) \quad (18)$$

Pour déterminer  $i(t)$  à partir des mesures de  $I(t)$  et  $E(t)$ , plusieurs méthodes d'analyse des déclins de fluorescence existent. Pour nos travaux, nous avons choisi de traiter les mesures de temps de vie par une méthode basée sur la statistique des moindres carrés (algorithme de Levenberg-Marquardt) grâce à un programme "maison" élaboré par G. Krishnamoorthy. L'utilisation du logiciel nécessite le binage des données d'un facteur 2 afin de travailler sur 2048 canaux.

L'analyse suppose que la loi de fluorescence corresponde à un modèle multi-exponentiel dans lequel l'intensité décline telle une somme de déclins individuels monoexponentiels :

$$I(t) = \sum \alpha_i \exp(-t / \tau_i) \quad (19)$$

où  $I(t)$  est l'intensité de fluorescence au temps  $t$  et  $\tau_i$  est le temps de vie individuel d'amplitude correspondante  $\alpha_i$ , tel que  $\sum \alpha_i = 1$ . L'analyse du déclin de fluorescence est considérée comme satisfaisante lorsque la valeur de  $\chi^2$  tend vers 1 et que la courbe des résidus est distribuée de façon aléatoire autour de 0. Le temps de vie moyen  $\langle \tau \rangle$  de fluorescence est calculé à partir des temps de vie de fluorescence individuels ( $\tau_i$ ) et de leurs amplitudes relatives ( $\alpha_i$ ) selon  $\langle \tau \rangle = \sum \alpha_i \tau_i$ .

Généralement, les fluorophores existent dans plusieurs états conformationnels au sein des protéines et présentent un temps de déclin  $\tau$  différent pour chaque conformation. Ainsi, chaque temps de vie  $\tau_i$  est supposé représenter un état conformationnel différent avec  $\alpha_i$  correspondant à la fraction de molécules présentes dans cette conformation.

Dans le cas des oligonucléotides marqués à la 2-Ap, le déclin de fluorescence nécessite quatre composantes exponentielles pour être correctement décrit. Pour réconcilier la différence de variation des temps de vie par rapport à celle du rendement quantique entre la

sonde libre et la sonde insérée dans un oligonucléotide, il convient d'ajouter à ces quatre conformations (ou familles de conformations) une conformation (famille) supplémentaire dite "noire". Celle-ci correspond à des conformations fortement "stackées" associées à des émissions de fluorescence plus rapides que la limite temporelle de détection de notre montage expérimental ( $\sim 40$  ps). La population  $\alpha_0$  d'espèces noires de la 2-Ap est calculée selon l'équation :

$$\alpha_0 = 1 - \frac{\tau_{free}}{\tau_{ODN} \times R_m} \quad (20)$$

où  $\tau_{free}$  est le temps de vie de la 2-Ap libre,  $\tau_{ODN}$  le temps de vie moyen mesuré de la 2-Ap au sein de l'oligonucléotide (en absence ou en présence de protéine), et  $R_m$  le rapport des intensités de fluorescence correspondantes ( $R_m = I_{free}/I$ ). Les amplitudes restantes  $\alpha_{ic}$  sont recalculées à partir des amplitudes mesurées d'après :  $\alpha_{ic} = \alpha_i \times (1 - \alpha_0)$ .

### 3.3. Mesure des déclin d'anisotropie de fluorescence

L'information obtenue par l'anisotropie résolue en temps est plus riche que la valeur d'anisotropie moyenne obtenue à l'état stationnaire. En effet, la forme que prend le déclin d'anisotropie résolue en temps dépend de la taille, de la forme, et de la flexibilité des molécules marquées, ce qui se traduit généralement par un déclin multi-exponentiel. Pour une molécule dont la forme dévie fortement d'une sphère par exemple, les temps de corrélation observés sont déterminés par les vitesses de rotation de la molécule selon ses différents axes, ce qui se traduit par un déclin multi-exponentiel. De même, les mouvements locaux (comme le mouvement de la sonde vis-à-vis de la molécule à laquelle elle est liée) ou segmentaux, induisent des composantes exponentielles supplémentaires. Enfin, il existe des cas de figure où la rotation des molécules est dite gênée (la molécule ne diffuse pas librement). Ce dernier cas de figure induit des déclin d'anisotropie présentant une anisotropie résiduelle non-nulle.

Le dispositif de mesure est identique à celui utilisé pour les mesures de déclin de fluorescence (Figure 35), à quelques points près. Le plan de polarisation du faisceau d'excitation (P) est alternativement vertical ou horizontal. Le polariseur d'émission est orienté verticalement. Les intensités de fluorescence I parallèle,  $I_{//}(t)$ , et I perpendiculaire,  $I_{\perp}(t)$ , sont

enregistrées alternativement toutes les 50 secondes, pendant des temps égaux, de façon à pouvoir comparer les intensités. Pour mesurer  $I_{\perp}(t)$ , on intercale une lame de quartz cristallin rotatoire (LR), dont l'épaisseur a été calculée selon la longueur d'onde d'excitation. Afin d'avoir des intensités d'excitation identiques, une lame de silice fondue (sans pouvoir rotatoire) est intercalée dans le faisceau d'excitation pour la mesure de  $I_{//}(t)$ . Ainsi le facteur correctif  $G$  appliqué à  $I_{\perp}(t)$  est-il égal à 1.

L'anisotropie de fluorescence résolue en temps,  $r(t)$ , est définie par :

$$r(t) = \frac{I_{//}(t) - G \times I_{\perp}(t)}{I_{//}(t) + 2G \times I_{\perp}(t)} \quad (21)$$

où  $I_{//}(t)$  et  $I_{\perp}(t)$  sont respectivement les réponses impulsionnelles des composantes polarisées parallèlement et perpendiculairement à la direction de polarisation de la lumière incidente,  $G$  est le facteur correctif de l'appareil, et  $I(t)$  est l'intensité de fluorescence totale.

Les déclin des composantes parallèle et perpendiculaire de l'émission sont données par :

$$I_{//}(t) = I(t) \frac{[1 + 2r(t)]}{3} \quad (22)$$

$$I_{\perp}(t) = I(t) \frac{[1 - r(t)]}{3} \quad (23)$$

La technique de déconvolution citée précédemment pour les durées de vie de fluorescence est appliquée à  $I_{//}(t)$  et  $I_{\perp}(t)$  en deux étapes. Dans un premier temps, les durées de vie,  $\tau_i$ , sont déterminées à partir de  $I(t)$  total. Puis, connaissant les  $\tau_i$ , on détermine les temps de corrélation de rotation,  $\phi_i$ , à partir de  $I_{//}(t)$  et  $I_{\perp}(t)$ .

L'anisotropie de fluorescence,  $r(t)$ , à chaque temps  $t$  suit une loi de décroissance exponentielle caractérisée par le(s) temps de corrélation de rotation,  $\phi_i$  :

$$r(t) = r_0 \sum_i \beta_i e^{-\frac{t}{\phi_i}} \quad (24)$$

où  $r_0$  est l'anisotropie fondamentale, et  $\beta_i$  correspond à la fraction de population associée au temps de corrélation de rotation  $\phi_i$ . Pendant l'analyse, les valeurs de  $\tau_i$  et  $\alpha_i$  sont fixées pour réduire le nombre de paramètres flottants. De même que pour les déclin de

fluorescence, l'analyse du déclin d'anisotropie est considérée comme satisfaisante lorsque la valeur de  $\chi^2$  tend vers 1 et que la courbe des résidus est distribuée de façon aléatoire autour de 0.

### 3.4. Conditions expérimentales

Taux de répétition des impulsions d'excitation : 4 MHz

Linéarité temporelle : 12,20 ps pour une acquisition dans 4096 canaux

Nombre total de coups par déclin :  $1 \times 10^6$

Taux de comptage inférieur à 7 kcps

Fentes du photomultiplicateur d'émission : 16 nm

Pour chaque échantillon, les mesures sont répétées au moins deux fois.

Pour les mesures de temps de vie de fluorescence des résidus Trp de la protéine SRA (408-643), les mesures sont effectuées sur 0,4  $\mu\text{M}$  de protéine. La longueur d'onde d'excitation utilisée est de 295 nm pour exciter sélectivement le Trp. La longueur d'onde d'émission est de 350 nm.

Pour les mesures de temps de vie et d'anisotropie résolue en temps réalisées sur les oligonucléotides marqués à la 2-Ap en absence ou en présence de protéine, une longueur d'onde d'excitation de 315 nm permet d'exciter sélectivement la 2-Ap. La longueur d'onde d'émission est de 370 nm. Les mesures sont effectuées sur 3  $\mu\text{M}$  d'oligonucléotides en absence ou en présence de 4  $\mu\text{M}$  de protéine.



## IV- RESULTATS ET DISCUSSION

- A. Publication 1 : "Site-specific characterization of the dynamics of UHRF1 SRA domain bound to 2-aminopurine labeled hemi-methylated oligonucleotides"

**Site-specific characterization of the dynamics of UHRF1 SRA domain bound to 2-aminopurine labeled hemi-methylated duplex oligonucleotides.**

Vanille J. Greiner<sup>1</sup>, Lesia Kovalenko<sup>2</sup>, Nicolas Humbert<sup>1</sup>, Ludovic Richert<sup>1</sup>, Catherine Birck<sup>3</sup>, Marc Ruff<sup>3</sup>, Olga A. Zaporozhets<sup>2</sup>, Christian Bronner<sup>1</sup>, Yves Mély<sup>1\*</sup>

<sup>1</sup> *Laboratoire de Biophotonique et Pharmacologie, UMR 7213 CNRS, Université de Strasbourg, Faculté de pharmacie, 74 route du Rhin, 67401 Illkirch, France*

<sup>2</sup> *Analytical Chemistry Department, Taras Shevchenko National University, 64 Volodymyrska Street, 01033 Kyiv, Ukraine*

<sup>3</sup> *Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U964 CNRS UMR 7104, Université de Strasbourg, Illkirch, France*

\* *Corresponding author*

## ABSTRACT

The protein UHRF1 plays a central role in the maintenance and transmission of epigenetic modifications. During the replication process, this protein recruits Dnmt1 to hemi-methylated CpG sites via its SET and RING Associated (SRA) domain, promoting the duplication of the methylation profiles. Tridimensional structure of the SRA/DNA complex revealed that the protein induces a base-flipping of the methyl-cytosine that enables a specific anchoring of the protein to hemi-methylated sites facilitating the recruitment of Dnmt1 to these strategic positions. In order to further understand the recognition mechanism of UHRF1, duplex oligonucleotides were labeled by 2-aminopurine, a fluorescent nucleoside analogue sensitive to environment, at various positions close to the single hemi-methylated CpG recognition site. Steady-state and time-resolved fluorescence spectroscopy measurements of these duplexes bound to the SRA domain of UHRF1 enabled us to site-specifically characterize the conformational changes induced by the binding of this domain. In agreement with the tridimensional structure of the SRA/DNA complex, our data suggest that the SRA domain is able to flip the methyl-cytosine while preserving the structure of the surrounding bases in the duplex. Remarkably, the SRA domain seems to bind with the same mechanism to hemi-methylated, fully-methylated and non-methylated duplexes, suggesting the protein plays a role of “reader” that scans the DNA sequence for hemi-methylated sites.

## INTRODUCTION

The fate of a cell is under the dependence of epigenetic mechanisms that regulate the expression of genes without affecting the DNA sequence. One of the main epigenetic modifications is DNA methylation, which patterns are set during development and constitute the "cell memory". Indeed, one of the first steps of the transmission of a cell to its descendants involves the recognition of the hemi-methylated CpG sites generated after DNA replication, and the duplication of the methylation profiles on the newly synthesized strand, mediated by the maintenance methyltransferase Dnmt1 (1,2). DNA methylation allows the regulation of gene expression by inhibiting the binding of transcriptional factors to their promoters or inducing the recruitment of chromatin remodeling complexes to specific regions of the genome (3). By being involved in various physiological processes like cellular differentiation and development, X chromosome inactivation and genomic imprinting (4), DNA methylation constitutes a key element in the epigenetic regulation.

The protein UHRF1 appears critical in the epigenetic code transmission, being at the interface between the modifications of DNA and histones. Indeed, UHRF1 possesses a SET and RING-associated (SRA) domain that recognizes hemi-methylated DNA (5-13), and the Tandem Tudor Domain (TTD) and Plant Homeo Domain (PHD) that can read histone modifications (14-16). By recruiting various actors of the epigenetic regulation to specific chromatin regions, UHRF1 is able to maintain and transmit epigenetic information (17-19). During the replication process, the UHRF1 protein likely contributes to the duplication of DNA methylation profiles by loading the Dnmt1 to hemi-methylated CpG sites, which enables the proper methylation of the target cytosine on the complementary strand (7,12). In line with the function of UHRF1, the SRA domain was shown to preferentially bind hemi-methylated DNA over non-methylated and fully-methylated DNA (5-11). The crystallographic structure of SRA in complex with hemi-methylated DNA revealed that this domain behaves like a hand in which the fingers flip out the methyl-cytosine that accommodates exactly the binding pocket of the protein (5,6,9). The minor groove loop, referred as the thumb, promotes the methyl-cytosine base-flipping, while the major groove loop, referred as the NKR finger due to its Asn-Lys-Arg sequence located at its tip, is involved in the reading and stabilization of the other three bases of the CpG duplex. This base-flipping mechanism is non-enzymatic and confers to UHRF1, the ability to anchor the

DNA at strategic positions facilitating the recruitment of Dnmt1 to hemi-methylated CpG sites.

To further understand the role of the methyl-cytosine base-flipping and the mechanism of interaction of the SRA domain with hemi-methylated DNA, we used duplex oligonucleotides (ODNs) labeled with 2-aminopurine (2-Ap) and investigated their dynamic changes in response to SRA binding. The 2-Ap is a fluorescent analogue of adenine that can be incorporated within DNA to substitute a natural base, playing the role of fluorescent reporter (20). This probe is highly sensitive to its micro-environment and especially to stacking with neighbor bases, giving thus local information within a nucleic acid (21-26). In our study, the introduction of 2-Ap at various positions within duplexes exhibiting a single hemi-methylated CpG site, allowed us to site-specifically investigate the conformational changes accompanying SRA binding and to correlate them with the known tridimensional structure of the SRA/DNA complex (5,6,9). We found that the SRA domain induces the base-flipping of the methyl-cytosine while preserving the global conformation of the duplex. Thus, in agreement with the crystallographic structure of the complex, the NKR finger of SRA likely occupies the free space left by the flipped methyl-cytosine and interacts with the surrounding bases. The duplex flexibility was particularly decreased on the 3' side of the methylated cytosine, suggesting that anchoring of the SRA domain to this strategic position induces conformational changes that in turn could promote the binding of the Dnmt1 to the cytosine to be methylated on the complementary strand. Finally, the SRA domain was shown to bind similarly to hemi-methylated, non-methylated and fully-methylated DNA, supporting a role of “reader” of the protein UHRF1 in which the NKR finger likely constitutes a scanner head for hemi-methylated CpG sites within DNA sequences.

## MATERIALS AND METHODS

### Materials

The SRA domain of hUHRF1 (SRA; residues 408-643) was expressed in *Escherichia coli* BL21-pLysS(DE3) and purified as described previously (27). Unmodified or labeled oligonucleotides (ODNs) were synthesized and HPLC-purified by IBA GmbH Nucleic Acids Product Supply (Germany). The sequence of the reference 12-bp duplex ODN was 5'-GGGCCXGCAGGG-3'/5'-CCCTGYYGGGCCC-3' with a single CpG site that was either unmethylated ( $X = Y = C$ ), hemi-methylated ( $X = 5mC$ ,  $Y = C$ ), or fully methylated ( $X = Y = 5mC$ ). This sequence found in the RB1 gene promoter is identical to the one used by Avvakumov et al. for the structure determination of the SRA/DNA complex (6). A 2'-deoxyribose-2-aminopurine (2-Ap) substituting the corresponding natural base was selectively introduced at different positions (3, 5, 6, 7, 8, 9, 8', 7', 6' or 5') in the vicinity of the CpG site within the sequence. Complementary strands were mixed in equal molar amounts in 25 mM Tris-HCl pH 7.5, 150 mM NaCl buffer and annealed by heating to 90°C and then cooled to room temperature. Absorption spectra were recorded on a Cary 400 spectrophotometer. Extinction coefficients comprised between 86,500 and 135,020  $M^{-1}cm^{-1}$  were used to determine the concentrations of duplex ODNs at 260 nm. Both the diffusion and buffer absorption were subtracted from the spectra. All experiments were performed at 20°C in 25 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.5 mM PMSF, 2 mM TCEP, in the presence of 0.04 % PEG 20,000 to prevent protein adsorption on the cuvette walls (28).

### Steady-state fluorescence spectroscopy

Fluorescence emission spectra were recorded at 20°C on a FluoroMax-3 or FluoroLog spectrofluorimeter (Jobin Yvon) equipped with a thermostated cell compartment. Excitation wavelength was at 315 nm to excite selectively the 2-Ap. Spectra were corrected for screening effects, buffer fluorescence and lamp fluctuations. Quantum yields of the labeled duplexes in the absence or presence of the SRA protein were determined by taking free 2-Ap riboside as a reference (quantum yield = 0.68 (20)). Measurements were performed on 3  $\mu M$  of ODN duplexes in the absence or presence of 4  $\mu M$  SRA protein.

## Time-resolved fluorescence spectroscopy

Time-resolved fluorescence measurements were performed with the time-correlated, single-photon counting technique. Excitation pulses were generated by a pulse-picked frequency-tripled Ti-sapphire laser (Tsunami, Spectra Physics) pumped by a Millennia X laser (Spectra Physics). Excitation wavelength was set at 315 nm, with a repetition rate of 4 MHz. The fluorescence emission was collected through a polarizer set at magic angle and a 16 mm band-pass monochromator (Jobin Yvon) at 370 nm. The single-photon events were detected with a micro-channel plate photomultiplier (Hamamatsu) coupled to a pulse pre-amplifier HFAC (Becker-Hickl) and recorded on a SPC-130 board (Becker-Hickl). The instrumental response function (IRF) was recorded using a polished aluminium reflector, and its full-width at half-maximum was ~50 ps.

The mean lifetime  $\langle\tau\rangle$  was calculated from the individual fluorescence lifetimes ( $\tau_i$ ) and their relative amplitudes ( $\alpha_i$ ) according to  $\langle\tau\rangle = \sum \alpha_i \tau_i$ . The population,  $\alpha_0$ , of dark species of 2-Ap was calculated by:  $\alpha_0 = 1 - \tau_{\text{free}} / (\tau_{\text{ODN}} \times R_m)$ , where  $\tau_{\text{free}}$  is the lifetime of the free 2-Ap,  $\tau_{\text{ODN}}$  is the measured mean lifetime of 2-Ap within the ODN duplex and  $R_m$  is the ratio of their corresponding steady-state fluorescence intensities ( $R_m = I_{\text{free2Ap}}/I_{\text{ODN}}$ ). The remaining amplitudes,  $\alpha_{ic}$ , were recalculated from the measured amplitudes according to:  $\alpha_{ic} = \alpha_i \times (1 - \alpha_0)$ .

For time-resolved anisotropy measurements, the fluorescence decay curves were recorded at vertical and horizontal positions of the polarizer and analyzed by the following equations:

$$\begin{aligned}
 I_{\parallel}(t) &= I(t) \frac{[1 + 2r(t)]}{3} \\
 I_{\perp}(t) &= I(t) \frac{[1 - r(t)]}{3} \\
 r(t) &= \frac{I_{\parallel}(t) - G \times I_{\perp}(t)}{I_{\parallel}(t) + 2G \times I_{\perp}(t)} = r_0 \sum_i \beta_i e^{\frac{-t}{\tau_i}} \quad (1)
 \end{aligned}$$

where  $\beta_i$  are the amplitudes of the rotational correlation times  $\tau_i$ ,  $I_{\parallel}(t)$  and  $I_{\perp}(t)$  are the intensities collected at emission polarization parallel and perpendicular, respectively, to the polarization of the excitation beam,  $r_0$  the initial anisotropy value, and  $G$  is the geometry factor at the emission wavelength, determined in independent experiments. The theoretical



values of the rotational correlation times were calculated from the molecular masses ( $M$ ) of the molecules and their complexes, assuming spherical shapes, by:

$$\varphi = \frac{\eta M(v + h)}{RT} \quad (2)$$

where  $\eta$  is the viscosity (assumed to be 1 cP),  $T$  is the temperature (293 K),  $v$  is the specific volume of the particle (assumed to be 0.78 mL/g for proteins (29)),  $h$  is the hydration degree (assumed to be 0.4 mL/g for proteins) and  $R$  is the molar gas constant.

The cone semi-angle ( $\theta_0$ ) providing an estimate for the local motion of 2Ap modeled as a diffusion within a cone (30) was calculated as:

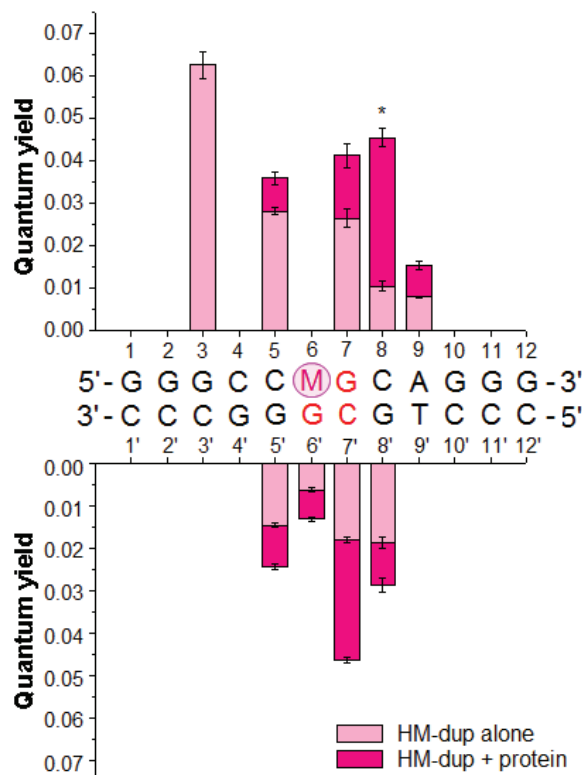
$$\theta_0 = \cos^{-1}(0.5 \times (\sqrt{1 + 8S} - 1)) \quad (3)$$

where  $S$  is the generalized order parameter:  $S = ((\beta'_2 + \beta'_3)/r_0)^{1/2}$ ,  $r_0$  is the initial anisotropy value, and  $\beta'_i$  are the amplitudes of the rotational correlation times, recalculated such that  $\sum \beta'_i = r_0$ .

Time-resolved intensity and anisotropy data were treated with a nonlinear least-square analysis using a homemade software (kindly provided by G. Krishnamoorthy). In all cases, the  $\chi^2$  values were close to 1, and the weighted residuals as well as the autocorrelation of the residuals were distributed randomly around zero, indicating an optimal fit.

## RESULTS

### Characterization of the 2-Ap labeled hemi-methylated duplex oligonucleotides



**Figure 1.** Quantum yield of the 2-Ap labeled hemi-methylated duplex ODNs in the absence or presence of SRA (408-643) domain. The sequence of the 12-bp ODN contains a single CpG recognition site (in red) with a methyl-cytosine at position 6 (M). The 2-Ap residue replaces the natural base at various positions (3, 5, 6, 7, 8, 9, 5', 6', 7', 8') within the ODN sequence. Experiments were performed with 3  $\mu$ M of 2-Ap labeled ODNs in the absence or presence of 4  $\mu$ M SRA domain in 25 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.5 mM PMSF, 2 mM TCEP, 0.04 % PEG 20,000 at 20°C. Excitation wavelength was 315 nm. The histogram values correspond to the average of at least two experiments. Standard deviations for these measurements are reported on the histogram.

\* Data obtained with a sequence where the G at position 8' facing the 2-Ap residue at position 8 was replaced by a T revealed an increased sensitivity of 2-Ap to the binding of SRA. This may be related to the fact that 2-Ap does not base pair efficiently with G8' (31), so that a larger population can be driven to an extra-helical conformation as a consequence of SRA binding (see explanation in the text and Tables S5 and S6 in the Supplementary data).

In order to site-specifically characterize the dynamics of the SRA domain in complex with the 12-bp hemi-methylated duplexes (HM-dup), the environment sensitive probe 2-Ap was introduced at various positions in the vicinity (HM5, HM5', HM6', HM7, HM7', HM8, HM8', HM9) or rather far (HM3) from the central CpG recognition site containing a hemi-methylated cytosine (Figure 1). As compared to the native duplex that exhibits a melting temperature of 55 °C, the melting temperatures of the 2-Ap labeled duplexes were 10 to 20

°C lower (data not shown), indicating that though the introduction of 2-Ap decreases the stability of the duplexes, their dissociation remains negligible at 20°C. In the labeled HM-dup, the 2-Ap emission is largely decreased as compared to free 2-Ap (20), with quantum yields being between 0.006 and 0.018 when the base flanking the 2-Ap residue is a guanine (Figure 1 and Table 1), the most efficient quencher of 2-Ap (32). The relatively high quantum yield value observed for HM3 (0.063) is probably due to the weaker level of stacking at the end of the duplex (33,34).

**Table 1.** Steady-state and time-resolved fluorescence parameters of 2-Ap-substituted hemi-methylated duplexes

	Q	$\alpha_0$	$\tau_1$ (ns)	$\alpha_1$	$\tau_2$ (ns)	$\alpha_2$	$\tau_3$ (ns)	$\alpha_3$	$\tau_4$ (ns)	$\alpha_4$	$\langle \tau \rangle$ (ns)
Free 2-Ap	0.680 <sup>a</sup>								10.2	1.00	10.2
HM3	0.063	0.28	0.05	0.52	0.51	0.06	3.12	0.06	8.82	0.08	1.32
HM3 + SRA	0.049	0.29	0.05	0.52	0.56	0.06	3.07	0.07	8.44	0.05	1.04
HM5	0.028	0.51	0.12	0.35	0.48	0.08	3.01	0.02	9.32	0.03	0.91
HM5 + SRA	0.036	0.39	0.08	0.45	0.59	0.07	3.06	0.06	9.45	0.03	0.89
HM7	0.027	0.22	0.04	0.52	0.41	0.21	2.79	0.03	8.35	0.02	0.52
HM7 + SRA	0.041	0.27	0.07	0.47	0.53	0.14	2.90	0.09	8.33	0.03	0.86
HM8	0.011	0.63	0.05	0.31	0.50	0.03	3.17	0.02	9.38	0.01	0.45
HM8 + SRA	0.046	0.62	0.10	0.21	0.81	0.06	3.33	0.08	8.60	0.04	1.81
HM9	0.008	0.71	0.07	0.25	0.44	0.02	3.31	0.01	10.43	0.01	0.41
HM9 + SRA	0.016	0.71	0.07	0.22	0.52	0.03	2.98	0.03	10.18	0.01	0.80
HM8'	0.019	0.33	0.06	0.55	0.49	0.07	3.18	0.04	8.80	0.01	0.43
HM8' + SRA	0.029	0.14	0.08	0.67	0.46	0.10	2.83	0.07	9.23	0.02	0.50
HM7'	0.018	0.74	0.09	0.18	0.59	0.04	3.64	0.03	9.33	0.01	1.03
HM7' + SRA	0.047	0.62	0.15	0.17	0.84	0.08	3.18	0.09	8.69	0.04	1.86
HM6'	0.006	0.73	0.05	0.23	0.37	0.03	2.77	0.01	9.72	0.01	0.37
HM6' + SRA	0.013	0.73	0.06	0.20	0.49	0.03	3.10	0.04	9.62	0.01	0.73
HM5'	0.015	0.69	0.05	0.25	0.42	0.03	3.22	0.02	9.25	0.01	0.72
HM5' + SRA	0.024	0.56	0.06	0.31	0.58	0.06	2.92	0.06	9.21	0.02	0.83

Experiments were performed with 3  $\mu$ M of 2-Ap labeled hemi-methylated duplexes in the absence or presence of 4  $\mu$ M SRA domain in 25 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.5 mM PMSF, 2 mM TCEP, 0.04 % PEG 20,000 at 20°C. Excitation and emission wavelengths were 315 nm and 370 nm, respectively. Q corresponds to the quantum yield. The best fit to the time-resolved fluorescence decay was obtained using four lifetimes where  $\tau_i$  correspond to the fluorescence lifetimes,  $\alpha_i$  to the relative amplitudes, and  $\langle \tau \rangle$  to the mean lifetimes. The amplitude  $\alpha_0$  of the dark species, as well as the amplitudes of the various lifetimes were calculated as described in the Materials and Methods section. Standard deviations are lower than 10 % for the quantum yield, 25 % for the lifetimes, 25 % for the amplitudes and 15 % for the mean lifetimes, respectively.

<sup>a</sup> Data from (20)

The fluorescence decays of the 2-Ap labeled duplexes exhibit four lifetimes attributed to different conformational states (Table 1). Moreover, as the differences in the quantum yields of the free 2-Ap and 2-Ap labeled duplexes were higher than the differences of their

corresponding mean lifetimes, an additional 2-Ap population corresponding to dark species with a lifetime shorter than the detection limit of our equipment ( $< 0.03$  ns) (22,23) was inferred. This population of dark species ( $\alpha_0$ ) together with the population ( $\alpha_1$ ) associated with the 30-50 ps lifetime is thought to correspond to the most stacked conformations of 2-Ap (32). These conformations are predominant for all 2-Ap labeled duplexes and exceed 90 % when the neighboring base is a guanine. The fourth population ( $\alpha_4$ ) exhibits a lifetime ( $\tau_4$ ) ( $\sim 9.27$  ns) close to that of the free 2-Ap and likely corresponds to extra-helical conformations (32). Except for HM3, where the  $\alpha_4$  population represents 8 %, this population is negligible (1 to 3 %) for most of the duplexes, confirming that the stacking of 2-Ap with its neighbors is less pronounced at the end of the duplex (34).

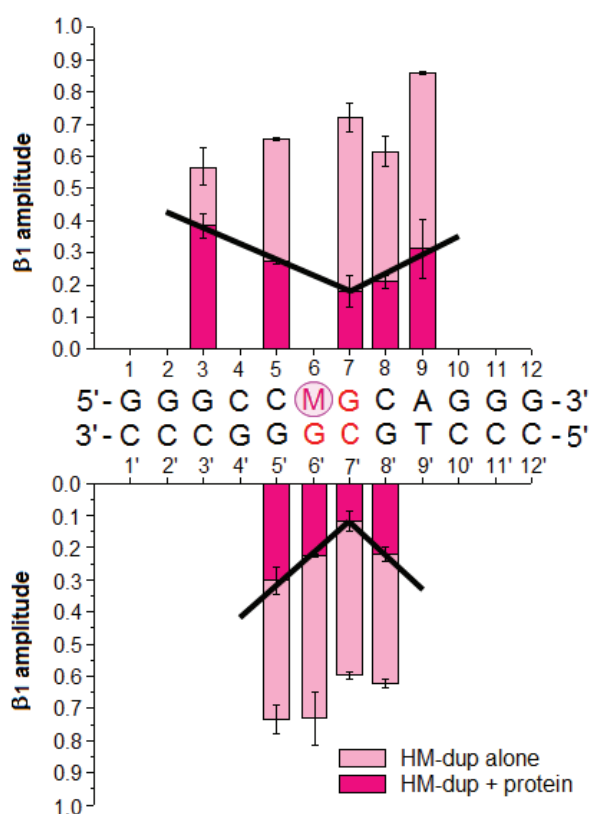
**Table 2.** Fluorescence anisotropy decay parameters of 2-Ap-substituted hemi-methylated duplexes

	$\varphi_1$ (ns)	$\beta_1$	$\varphi_2$ (ns)	$\beta_2$	$\varphi_3$ (ns)	$\beta_3$	S	$\theta_0$ (degree)
Free 2-Ap	0.08	1.00					0	90.0
HM3	0.37	0.56	2.47	0.44			0.660	41.1
HM3 + SRA	0.23	0.39	1.58	0.28	16.4	0.34	0.785	31.9
HM5	0.44	0.65	1.88	0.35			0.589	46.0
HM5 + SRA	0.24	0.27	2.64	0.32	17.0	0.41	0.853	26.0
HM7	0.61	0.72	2.38	0.28			0.528	50.0
HM7 + SRA	0.21	0.18	5.09	0.16	17.4	0.66	0.904	20.4
HM8	0.34	0.62	2.49	0.39			0.618	44.0
HM8 + SRA	0.26	0.21	2.23	0.21	17.2	0.58	0.888	22.5
HM9	0.22	0.86	3.15	0.14			0.374	60.1
HM9 + SRA	0.09	0.31	4.20	0.20	17.3	0.49	0.825	28.1
HM8'	0.38	0.62	2.68	0.38			0.615	44.2
HM8' + SRA	0.29	0.22	1.64	0.22	16.7	0.56	0.883	23.1
HM7'	0.25	0.60	1.83	0.40			0.636	42.8
HM7' + SRA	0.34	0.12	2.02	0.23	16.8	0.65	0.939	16.3
HM6'	0.21	0.73	1.64	0.27			0.505	51.5
HM6' + SRA	0.12	0.23	5.10	0.25	15.9	0.52	0.880	23.4
HM5'	0.19	0.73	1.74	0.27			0.514	50.9
HM5' + SRA	0.15	0.30	4.02	0.34	17.5	0.36	0.835	27.5

$\varphi_i$  are the rotational correlation times,  $\beta_i$  their amplitudes, S is the generalized order parameter and  $\theta_0$  the cone semi-angle for 2-Ap local motion calculated as described in the Materials and Methods section. Standard deviations are lower than 20 % for both the rotational correlation times and their amplitudes.

Time-resolved fluorescence anisotropy measurements were performed to further report on the 2-Ap dynamics. In all duplexes, the value of the initial anisotropy  $r_0$  of the 2-Ap residue was about 0.32 (data not shown), in line with 0.33 value of 2-Ap in 77 % glycerol

(35). Fluorescence anisotropy decays of the 2-Ap-labeled ODNs showed two correlation times  $\phi_1$  and  $\phi_2$  (Table 2). The shorter component  $\phi_1$  (0.19 - 0.61 ns) was attributed to the local rotation of the 2-Ap probe (32,34). The high proportion of the  $\phi_1$  population, representing up to 86 % (Figure 2), revealed the high mobility of the 2-Ap base, at least in its least stacked conformations that contribute most to the 2-Ap emission. This mobility can be further quantified by using the semi-angle cone model (30,32), which suggests that the 2-Ap base can explore angular ranges of 41 to 60°, depending on the labeled position. Moreover, the slow correlation time of 1.64 to 3.15 ns, was significantly shorter than the theoretical correlation time (3.8 ns) calculated for the motion of a sphere with the same molecular mass than the duplex (29). This indicates a significant contribution of segmental motions, indicative of the high flexibility of the duplexes (34).



**Figure 2.**  $\beta_1$  amplitudes of the rotational correlation times  $\phi_1$  of the 2-Ap labeled HM duplexes in the absence or the presence of the SRA (408-643) domain. The 12-bp ODN duplexes contain a single CpG recognition site (in red) with a methyl-cytosine at position 6 (M). The 2-Ap residue substitutes the natural base at various positions (3, 5, 6, 7, 8, 9, 5', 6', 7', 8') within the ODN sequence. Black lines highlight the strong restriction of the local motion of the 2-Ap residues induced by the binding of the protein. This restriction is the highest at the level of the CpG recognition site. Experiments were performed with 3  $\mu$ M of 2-Ap labeled ODNs in the absence or presence of 4  $\mu$ M SRA domain in 25 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.5 mM PMSF, 2 mM TCEP, 0.04 % PEG 20,000 at 20°C. Excitation and emission wavelengths were 315 nm and 370 nm, respectively. The histogram values correspond to the average of at least two experiments. Standard deviations are reported on the histogram.

## Effects of the SRA domain on the dynamics of the 2-Ap labeled hemi-methylated duplexes

In order to site-selectively investigate the interaction of the UHRF1 SRA domain with HM duplexes, fluorescence measurements were performed on 2-Ap labeled HM-dup in the presence of the SRA (408-643) domain. In agreement with previous studies (5-11), FRET experiments using Cy3-labeled SRA protein and Cy5-labeled duplexes, showed that the SRA domain of UHRF1 interacts with a strong affinity with HM duplexes ( $K_d = 0.08 \mu\text{M}$ , Supplementary data Figure S1). Moreover, competition experiments revealed that the substitution of a natural base by a 2-Ap residue only slightly decreased the SRA binding affinity (Supplementary data Figure S2). According to these data, all following experiments were performed in conditions where at least 80 % of the duplexes were complexed to SRA.

The binding of the SRA domain to the 2-Ap labeled HM duplexes increased the 2-Ap fluorescence quantum yield by 1.2- to 3.7-fold, depending on the labeled position (Figure 1, Table 1 and Figure 3). This change in quantum yield was accompanied by a 2 to 6 nm blue shift of its maximum emission wavelength (data not shown), indicating a slight decrease in the polarity of the 2-Ap environment. Time-resolved fluorescence intensity decays revealed that the SRA-induced increase of the 2-Ap quantum yield was mainly due to an increase of the populations of the least stacked 2-Ap conformers, as can be seen from the changes in the  $\alpha_3$  and  $\alpha_4$  values (Table 1 and Supplementary Figure S3). In line with the numerous contacts of SRA with the duplexes (5,6,9), these changes suggest that SRA restricts the stacking and collisions of 2-Ap with its neighbor residues. Time-resolved anisotropy decays further revealed that SRA strongly restricts the local motion of the 2-Ap bases (Table 2 and Figure 2), as could be seen from the decrease of  $\beta_1$  and  $\theta_0$  values. Moreover, the appearance of a third component  $\varphi_3$  (~ 17 ns) is fully consistent with the theoretical 16.8 ns value calculated for the tumbling of a 1:1 spherical complex, suggesting that the  $\varphi_3$  component only describes the overall motion of the complex with no contribution of segmental motions, which are now described by the  $\varphi_2$  component. Taken together, these data indicate that the binding of the SRA domain leads to a strong decrease of both the global flexibility of the duplexes and the local mobility of the bases.

Comparison of the data at the different labeled positions indicates that the strongest variations in the presence of SRA are observed at positions directly surrounding the flipped

methyl-cytosine. Indeed, dramatic decreases in the  $\beta_1$  amplitude values (from ~69 % to ~22 %) (Table 2 and Figure 2) were observed when the 2-Ap residue replaces the base opposite to the methyl-cytosine in the HM duplex (HM6') or one of the nucleotides of the base-pairs flanking the methyl-cytosine (HM5, HM5', HM7 and HM7'). These data indicate that SRA strongly restricts the local motion of the bases at these sites, in line with the 3D structure of the SRA/DNA complex showing direct contacts of these bases through van der Waals and/or hydrogen bonding interactions with the NKR finger (Asp 489, Lys 490, Arg 491) of the SRA domain (5,6,9). The somewhat larger changes observed when 2-Ap replaces G7 or C7' as compared to C5 or G5' (Figure 2) indicate a stronger anchoring of the SRA domain to the nucleotides downstream of the flipped cytosine (6,9). Surprisingly, in spite of the flip of the methyl-cytosine at position 6, the fluorescence quantum yield and lifetimes of 2-Ap at the surrounding positions (HM5, HM5', HM6', HM7 and HM7') were only moderately affected by the binding of SRA (Table 1, Figure 1 and Figure 3, and Supplementary Figure S3). This situation was previously observed in other studies where only one stacking neighbor was sufficient to induce a large fluorescence quenching of the 2-Ap (36-39). As a consequence, our data suggest that the NKR residues of SRA seem to fill the space left by the flipped methyl-cytosine, stabilizing the intra-helical environment of the surrounding bases (5,6,9).

In addition to confirming interactions that were previously observed in the 3D structure of the SRA/DNA complex, our data further revealed direct contacts between the SRA domain and the base-pair 8.8'. Indeed, we observed a dramatic decrease in the  $\beta_1$  amplitude on SRA binding to HM duplexes where the natural bases at position 8 or 8' were substituted by 2-Ap (Table 2 and Figure 2), indicating that SRA strongly restricts the local motion of 2-Ap at these two positions. As for the positions directly surrounding the methyl-cytosine, binding of SRA only modestly changed the values of the 2-Ap fluorescence quantum yield and lifetime parameters at position 8' (Table 1, Figure 1 and Figure 3, and Supplementary Figure S3), indicating that SRA also preserves the native structure of the duplex at this position. In contrast, a significant 3.7 fold increase of the 2-Ap quantum yield accompanied by a consequent redistribution of the fluorescence lifetimes was observed on binding of SRA to HM8. Since the base facing the 2-Ap residue at position 8 is a guanine, the strong changes in the fluorescence parameters of HM8 may be related to the unstable mismatch 2-Ap.G base-pair that could favor extra-helical conformations of 2-Ap. To check this hypothesis, the measurements were repeated with a sequence where the G facing the 2-Ap residue was substituted by a T (Supplementary data HM8T, Tables S5 and S6). This substitution did not



affect the SRA-induced changes in the time-resolved anisotropy parameters, but dramatically decreased the changes in the 2-Ap fluorescence quantum yield and lifetime parameters, which became comparable to those observed at position 8'. As a consequence, data with 2-Ap substitutions at positions 8 and 8' strongly suggest that SRA contacts both residues and stabilizes this base-pair in its native conformation in the duplex. Nevertheless, it cannot be excluded that the restriction of the local mobility and the stabilization of both bases indirectly result from the interaction of the SRA domain with the neighbor bases at positions 7 and 7'.

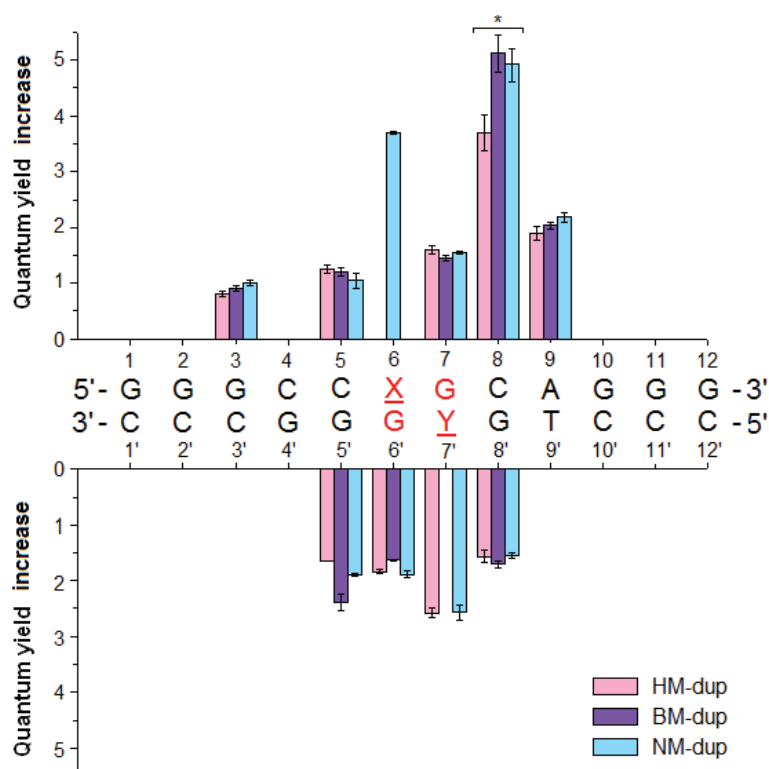
Since Hashimoto et al. also mentioned a base-flipping of the adenine at position 9 (9), we substituted this residue by 2-Ap (HM9) and measured the changes in its fluorescence parameters after binding of the SRA domain. The limited increase (1.9-fold) in the 2-Ap fluorescence quantum yield, as well as the moderate redistribution of its fluorescence lifetime parameters (Table 1, Figure 1 and Figure 3, and Supplementary Figure S3) are inconsistent with a base-flipping that generally increases the 2-Ap fluorescence by at least 10 times (38,40-47). These limited changes in the fluorescence parameters merely suggest a conformational change of the base at position 9 on SRA binding. In addition, the time-resolved anisotropy parameters further indicate that SRA restricts the local motion of the base at position 9 (Table 2 and Figure 2), likely through a direct contact with the protein or indirectly through the stabilized flanking bases.

In contrast to the aforementioned positions, substitution of the natural base at position 3 (HM3) led to no significant changes in the 2-Ap fluorescence quantum yield (Table 1, Figure 1 and Figure 3) and time-resolved parameters (Table 1, Table 2 and Figure 2, and Supplementary Figure S3) on SRA binding. This confirms that this base is located outside of the recognition site of the protein and that the SRA domain does not interact with this base.

All together the steady-state and time-resolved fluorescence changes observed with 2-Ap substitutions at different positions in the HM duplex confirm the presence of numerous contacts between the bases and the SRA residues that largely restrict the local mobility of the bases. This confirmation is of importance, since our data have been obtained in the more relevant context of diluted solutions. In addition, our data revealed that the SRA protein flips the methyl-cytosine base, while preserving the global structure of the duplex. This likely results from the ability of the NKR finger (Asp 489, Lys 490, Arg 491) to fit into the space left by the flipped methyl cytosine in the DNA helix, in order to interact with the adjacent bases and stabilize the intra-helical environment (5,6,9). The strong stabilization activity of the SRA domain close to the flipped methyl cytosine is confirmed by the fact that the

flexibility of the HM duplex is particularly decreased at positions at and close to the recognition site (Table 2 and Figure 2).

### Comparison of the SRA domain effects on the dynamics of the 2-Ap labeled non-methylated, hemi-methylated, or fully-methylated duplexes



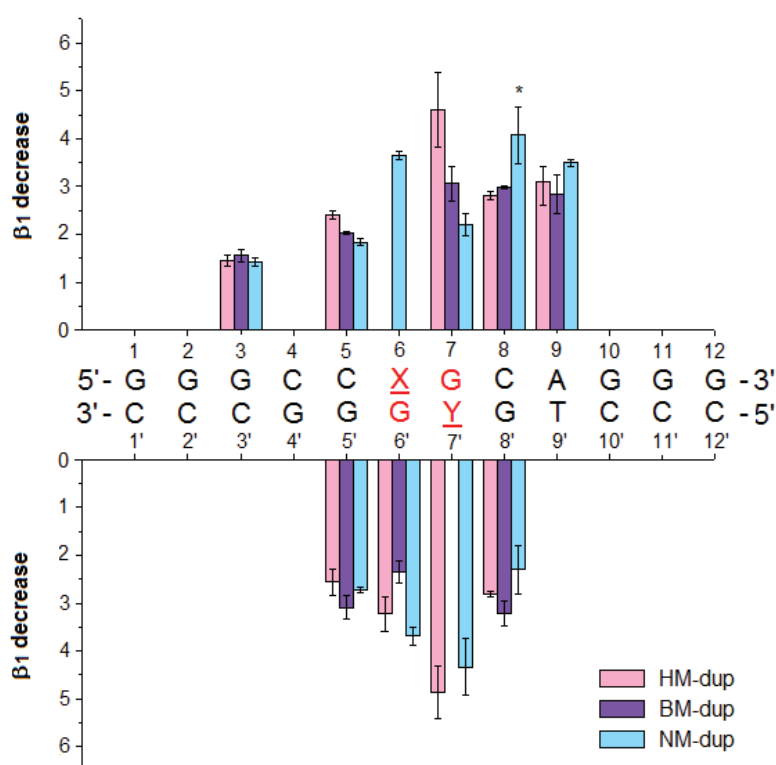
**Figure 3.** Changes of quantum yield of the 2-Ap labeled duplexes on addition of the SRA (408-643) domain. The sequence of the 12-bp ODN contains a single CpG recognition site (in red) that is either unmethylated (NM-dup: X = Y = C), hemi-methylated (HM-dup: X = 5mC, Y = C), or fully methylated (BM-dup: X = Y = 5mC). The 2-Ap residue replaces the natural base at various positions (3, 5, 6, 7, 8, 9, 5', 6', 7', 8') within the duplexes. Experiments were performed with 3  $\mu$ M of 2-Ap labeled duplexes in the absence or the presence of 4  $\mu$ M SRA domain in 25 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.5 mM PMSF, 2 mM TCEP, 0.04 % PEG 20,000 at 20°C. Excitation and emission wavelengths were 315 nm and 370 nm, respectively. For each position in the various duplexes, mean values  $\pm$  standard deviations are reported for at least two experiments.

\* Data obtained with a sequence where the G at position 8' facing the 2-Ap residue at position 8 was replaced by a T revealed an increased sensitivity of 2-Ap to the binding of SRA. This may be related to the fact that 2-Ap does not base pair efficiently with G8' (31), so that a larger population can be driven to an extra-helical conformation as a consequence of SRA binding (see explanation in the text and Tables S5 and S6 in the Supplementary data).

To further characterize the specific features of the interaction of the SRA domain with HM duplexes, we comparatively investigated with the same techniques the interaction of the SRA domain with the corresponding 2-Ap labeled duplexes, where the CpG recognition site was either non-methylated (NM) or methylated on both strands (BM). In agreement with previous studies (5-11), titrations by FRET confirmed that the SRA domain exhibits a

preferential affinity for HM duplexes, with a four-fold higher affinity as compared to NM- and BM duplexes ( $K_d \sim 0.28 \mu\text{M}$  and  $0.25 \mu\text{M}$  for NM and BM duplexes respectively, Supplementary data Figure S1).

In the absence of SRA, the higher values of the 2-Ap quantum yield for HM7 and BM7 as compared to NM7, and for BM6' as compared to HM6' and NM6' (Table 1 and Supplementary data Tables S1 and S2) revealed that methylated cytosine is a less efficient quencher of 2-Ap than the non-methylated cytosine.



**Figure 4.** Changes of the  $\beta_1$  amplitudes associated to the rotational correlation times  $\phi_1$  of the 2-Ap labeled duplexes on addition of the SRA (408-643) domain. The duplexes contains a single CpG recognition site (in red) that is either unmethylated (NM-dup: X = Y = C), hemi-methylated (HM-dup: X = 5mC, Y = C), or fully methylated (BM-dup: X = Y = 5mC). The 2-Ap residue substitutes the natural base at various positions (3, 5, 6, 7, 8, 9, 5', 6', 7', 8') within the duplexes. Experiments were performed with  $3 \mu\text{M}$  of 2-Ap labeled duplexes in the absence or the presence of  $4 \mu\text{M}$  SRA domain in 25 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.5 mM PMSF, 2 mM TCEP, 0.04 % PEG 20,000 at  $20^\circ\text{C}$ . Excitation and emission wavelengths were 315 nm and 370 nm, respectively. For each position in the various duplexes, mean values  $\pm$  standard deviations are reported for at least two experiments.

\* The data obtained with a sequence where the G facing the 2-Ap residue was replaced by a T revealed an increased sensitivity of 2-Ap at position 8 to the binding of SRA only for NM duplexes (Supplementary data Tables S5 and S6).

Remarkably, despite the preferential binding of the SRA domain to HM duplexes, and the flip of the methyl-cytosine thought to be specifically induced by the SRA domain on HM duplexes, the fluorescence parameters were found remarkably similar for all three types of duplexes at the various labeled positions (Figures 3 and 4, Tables 1 and 2, and

Supplementary data Tables S1, S2, S3 and S4). This suggests that the SRA domain forms similar complexes with HM-, NM- and BM duplexes. In fact, only slight differences in the fluorescence parameters are observed at positions at or close to the recognition site. When the 2-Ap residue substitutes the G6' base, opposite to the target cytosine, the changes in fluorescence quantum yield and the lifetime amplitudes are lower for the BM6' duplex as compared to the NM6'- and HM6' duplexes (Figure 3 and Supplementary data Figure S3, Table 1 and Supplementary data Tables S1 and S2). Moreover, comparison of the  $\beta_1$  amplitudes indicates that the local motion of 2-Ap is less restricted for the BM6' duplex than for the two other types of duplexes (Figure 4, Table 2 and Supplementary data Tables S3 and S4), suggesting that methylation of the neighbor C7' residue lowers the interactions with the NKR finger, probably as a consequence of a steric clash induced by the additional methyl group (5,6,9). The lower decrease of the  $\beta_1$  amplitude observed for BM7 as compared to HM7 could also be a consequence of this steric clash that likely decreases the interaction of the G7 base with the NKR finger. In sharp contrast, at position 5', SRA was found to induce more conformational changes on binding to the BM5' duplex as compared to the HM5'- and NM5' duplexes. Indeed, a larger increase (2.4 fold) of the 2-Ap quantum yield is observed when the protein binds to BM5' as compared NM5' and HM5' (1.9- and 1.6-fold, respectively) (Figure 3, Table 1 and Supplementary data Tables S1 and S2). This larger increase in quantum yield is mostly due to the more pronounced increase in the ( $\alpha_3 + \alpha_4$ ) populations (Supplementary data Figure S3, Table 1 and Supplementary data Tables S1 and S2). Moreover, the larger decrease in the  $\beta_1$  amplitude observed on binding of SRA to the BM5' duplex as compared to the NM5'- and HM5'- duplexes (Figure 4, Table 2 and Supplementary data Tables S3 and S4), suggest a stronger interaction of the base at position 5' with the NKR finger in the complex of SRA with the BM duplex. Similar SRA-induced differences in the  $\beta_1$  amplitude were also observed at position 5 in duplexes where the G5' residue was substituted by a T residue to prevent the destabilization due to the mismatch of G with 2-Ap (Supplementary data Tables S5 and S6). Thus, in complexes of the SRA domain with BM duplexes, the position of the NKR finger seems to be shifted, likely as a consequence of the steric hindrance with the methyl group of the C7' residue. This shift of the NKR finger weakens the interaction with the G6' residue, but strengthens the interaction with the base-pair 5.5'. Another difference is observed at position 7 where the  $\beta_1$  amplitude indicates that the local motion of the G7 residue is less restricted in the NM7 duplex than in the HM7- and BM7 duplexes after binding of the SRA domain (Figure 4, Table 2 and

Supplementary data Tables S3 and S4). Thus, in the absence of the methylated-cytosine, the SRA domain binds less avidly to this strategic position, probably as the consequence of the inability of the NKR finger to flip the non methylated cytosine at position 6 and thus, to insert into the duplex. This reduced interaction of the NKR finger with the G7 residue may explain in part the lower affinity of NM7 as compared to HM7 for the SRA domain.

For the NM duplex, fluorescence measurements were also performed with the 2-Ap residue substituting the non-methylated cytosine at position 6 within the CpG recognition site of the SRA domain (NM6). As expected, the data with this NM6 duplex revealed that this position is highly destabilized by the SRA domain. Indeed, we observed a 3.7-fold increase of the 2-Ap fluorescence quantum yield, associated with a 4-fold increase of the least-stacked 2-Ap populations (Figure 3, and Supplementary data Figure S3 and Table S1). These changes are well below those expected if the C6 base would be flipped out of the duplex (38,40-47), but nevertheless they indicate that the base is significantly moved from its natural position in the duplex. Moreover, the large drop in the  $\beta_1$  amplitude (from 67 % to 19 %) shows that the local motion of the C6 base is largely restricted by the binding of SRA, suggesting that the C6 base is firmly bound to SRA residues (Figure 4 and Supplementary data Table S3). Very similar changes in the fluorescence parameters were obtained when the guanine base facing the 2-Ap residue was replaced by a thymine within the NM duplex (Supplementary data Tables S5 and S6), indicating that the observed changes are independent on the opposite base. Thus, our data suggest a strong and direct interaction of the SRA domain with the base at position 6 within the SRA/NM complex. Due to the overall similarities in the binding of SRA to HM- and NM duplexes, and the significant perturbations observed at this strategic position in the NM duplex, it can be inferred that the SRA domain interacts with C6 through its thumb and its NKR finger and partly moves it out of the duplex. It should be noticed that if a full flipping of this base would occur, the 2-Ap residue would be sandwiched between Tyr 478 and Tyr 466 residues, resulting in  $\pi$ -stacking interactions (5,6,9), which could possibly quench the 2-Ap fluorescence (44,48,49) and thus mask the observation of the base-flipping. Thus, the real extent of the flipping of the residue at position 6 in the NM duplex remains elusive.

Taken together, our data indicate that the SRA domain seems to bind HM-, BM- and NM duplexes in a similar way.

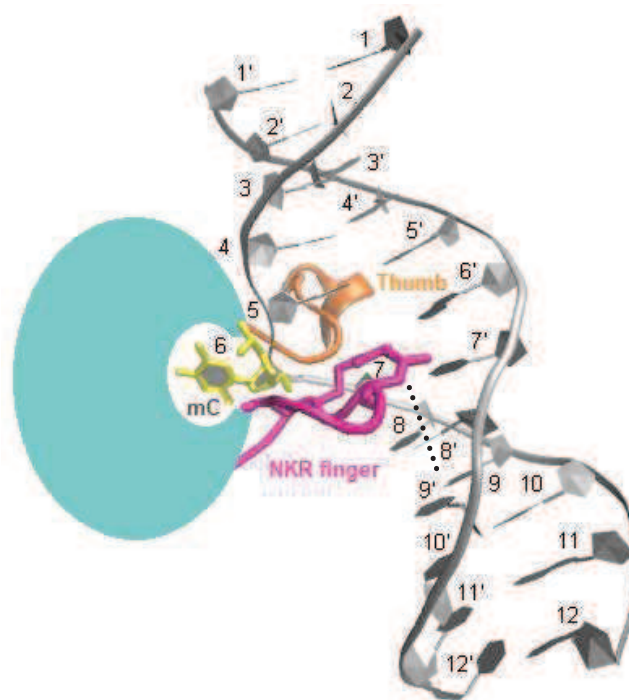
## DISCUSSION

In this study, we investigated the interaction of the UHRF1 SRA domain with 2-Ap substituted duplexes of 12-bp containing a single CpG recognition site that was either non-methylated, hemi-methylated or fully-methylated. Using steady-state and time-resolved fluorescence spectroscopy, we characterized site-specifically the dynamics of the interaction of the SRA domain with hemi-methylated DNA and compared it with non-methylated and bi-methylated DNA.

In the absence of protein, the duplexes show highly flexible structures (Table 2 and Supplementary data Tables S3 and S4), in which the 2-Ap residue adopts multiple conformations with a predominance of the most stacked conformers (Table 1). The quantum yield of the various 2-Ap labeled duplexes confirmed that the guanine residue is the most efficient quencher of the 2-Ap and revealed that methyl-cytosine is a less efficient quencher than the non-modified cytosine (Table 1 and Supplementary data Tables S1 and S2). Binding of the SRA domain was found to constrain both the overall flexibility of the duplexes and the local mobility of the bases, in line with the numerous contacts of SRA residues with the duplexes. Strong similarities were observed for the interaction of the SRA domain with non-methylated, hemi-methylated and bi-methylated duplexes, suggesting that the protein adopts similar binding modes with the different types of duplexes (Figures 3 and 4, and Supplementary data Figure S3). Thus, even if a full base flipping may not occur when the SRA domain binds to NM or BM duplexes, the protein seems to insert its NKR finger within the double-stranded DNA in the same way as in the complexes with HM duplexes. Moreover, as in NM duplexes, the protein shows a stronger tendency to perturb the base at position 6 over the opposite base in the palindromic CpG site, the SRA domain appears somewhat sensitive to the sequence of the bases surrounding its recognition site. Importantly, the absence of drastic changes in the 2-Ap fluorescence parameters when the target cytosine is flipped-out by the SRA domain suggests that the protein maintains the structure of the duplex during the flipping process. In line with the 3D structure of the SRA/DNA complex, this may result from the ability of the NKR finger (Asp 489, Lys 490, Arg 491) to occupy the space left by the flipped cytosine within the DNA helix and to stabilize the surrounding bases through direct contacts that mimic the natural intrahelical environment of duplexes (Figure 5) (5,6,9). In addition to the expected interactions of the SRA domain with the bases 5, 5', 6', 7 and 7' that directly surround the methyl-cytosine, our data indicate SRA-induced



conformational changes for bases at more distant positions, such as the 8.8' base-pair and the residue at position 9 (Tables 1 and 2). These changes could be due to direct contacts with the SRA domain or indirect changes in the duplex properties through the interaction of the SRA domain with the CpG recognition site. Interestingly, our data further indicate a preferential interaction of the SRA domain with the bases downstream of the methyl-cytosine and notably, with the G7.C7' base-pair (Figure 2). This preferential interaction and the accompanying conformational changes at this strategic position could in turn actively promote the recruitment of Dnmt1 at this site in order to enable the methylation of the cytosine residue at the 7' position opposite to the methyl-cytosine in the CpG recognition site.



**Figure 5.** Schematic representation of the SRA domain bound to hemi-methylated DNA (Based on Protein Data Bank code 3CLZ). The DNA duplex is shown in gray with the flipped methylated deoxycytidine (mC) in yellow. The body of the SRA domain is represented in blue with a binding pocket able to fit the methyl-cytosine. The thumb (orange) is thought to promote the flip of mC whereas the NKR finger (magenta) fills the resulting hole within the DNA helix, stabilizing the surrounding bases by direct contacts. The Asn 489 and Arg 491 residues of the NKR finger, that take part in most of these contacts, are shown in stick format. Dotted line represents the potential interactions between the SRA domain and the bases at positions 8, 8' and 9.

Despite the high affinity of the SRA domain for hemi-methylated DNA, the protein exhibits a modest preference for this type of DNA over non-methylated and bi-methylated DNA (8,10,11). Moreover, the SRA domain can even bind with a significant affinity to DNA duplexes lacking the CpG recognition site (8) (Supplementary data Figure S1). These observations are consistent with a “reader” role of the SRA domain that, through its



significant affinity for non-specific DNA, could slide along a DNA duplex to scan for hemi-methylated CpG sites (17,18). Since the binding mode of the protein seems similar for the various duplexes (Figures 3 and 4, and Supplementary data Figure S3), the thumb and the NKR finger of the SRA domain likely act together to scan the DNA sequence in search of hemi-methylated sites. The thumb is thought to interact with the DNA on the minor groove where it can promote the flipping of the bases, while the NKR finger is inserted from the opposite direction in the major groove and can thus act as a read head that is able to differentiate the methylation state of the strands (5,6,9). These two elements probably work together to recognize the hemi-methylated sequences. For the non-methylated DNA, unfavorable interaction energy likely favors the unbound state of the non-methylated DNA (50), preventing the anchoring of the protein to non-methylated CpG sites (as seen from the high overall flexibility kept for the NM7/SRA complexes). For the bi-methylated DNA sequences, the methyl groups on both cytosines in the CpG recognition sites, seem to cause a steric clash that destabilizes the conformation of the NKR finger within the DNA (as seen from the results obtained for BiM5' and BiM6'), accounting for the decrease in binding to fully-methylated DNA (5,6,9). Thus, the exact fitting of the methyl-cytosine within the binding pocket of the SRA domain (5,6,9), the favorable interaction energy of the bound hemi-methylated cytosine over its unbound state (50), and the stabilization of the surrounding bases by the NKR finger insertion (5,6,9), probably allow the complete base-flipping of this base and explain the preferential anchoring of the protein to hemi-methylated CpG sites.

## REFERENCES

1. Jones, P.A. and Liang, G. (2009) Rethinking how DNA methylation patterns are maintained. *Nat Rev Genet*, **10**, 805-811.
2. Jurkowska, R.Z., Jurkowski, T.P. and Jeltsch, A. (2011) Structure and function of mammalian DNA methyltransferases. *Chembiochem*, **12**, 206-222.
3. Rottach, A., Leonhardt, H. and Spada, F. (2009) DNA methylation-mediated epigenetic control. *Journal of cellular biochemistry*, **108**, 43-51.
4. Smallwood, S.A. and Kelsey, G. (2012) De novo DNA methylation: a germ cell perspective. *Trends Genet*, **28**, 33-42.
5. Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y. and Shirakawa, M. (2008) Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature*, **455**, 818-821.
6. Avvakumov, G.V., Walker, J.R., Xue, S., Li, Y., Duan, S., Bronner, C., Arrowsmith, C.H. and Dhe-Paganon, S. (2008) Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. *Nature*, **455**, 822-825.

7. Bostick, M., Kim, J.K., Esteve, P.O., Clark, A., Pradhan, S. and Jacobsen, S.E. (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science (New York, N.Y.)*, **317**, 1760-1764.
8. Frauer, C., Hoffmann, T., Bultmann, S., Casa, V., Cardoso, M.C., Antes, I. and Leonhardt, H. (2011) Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *PloS one*, **6**, e21306.
9. Hashimoto, H., Horton, J.R., Zhang, X., Bostick, M., Jacobsen, S.E. and Cheng, X. (2008) The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature*, **455**, 826-829.
10. Qian, C., Li, S., Jakoncic, J., Zeng, L., Walsh, M.J. and Zhou, M.M. (2008) Structure and hemimethylated CpG binding of the SRA domain from human UHRF1. *The Journal of biological chemistry*, **283**, 34490-34494.
11. Rottach, A., Frauer, C., Pichler, G., Bonapace, I.M., Spada, F. and Leonhardt, H. (2010) The multi-domain protein Np95 connects DNA methylation and histone modification. *Nucleic acids research*, **38**, 1796-1804.
12. Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K. *et al.* (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature*, **450**, 908-912.
13. Unoki, M., Nishidate, T. and Nakamura, Y. (2004) ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. *Oncogene*, **23**, 7601-7610.
14. Arita, K., Isogai, S., Oda, T., Unoki, M., Sugita, K., Sekiyama, N., Kuwata, K., Hamamoto, R., Tochio, H., Sato, M. *et al.* (2012) Recognition of modification status on a histone H3 tail by linked histone reader modules of the epigenetic regulator UHRF1. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 12950-12955.
15. Nady, N., Lemak, A., Walker, J.R., Avvakumov, G.V., Kareta, M.S., Achour, M., Xue, S., Duan, S., Allali-Hassani, A., Zuo, X. *et al.* (2011) Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein. *The Journal of biological chemistry*, **286**, 24300-24311.
16. Xie, S., Jakoncic, J. and Qian, C. (2012) UHRF1 double tudor domain and the adjacent PHD finger act together to recognize K9me3-containing histone H3 tail. *Journal of molecular biology*, **415**, 318-328.
17. Bronner, C., Fuhrmann, G., Chedin, F.L., Macaluso, M. and Dhe-Paganon, S. (2010) UHRF1 Links the Histone code and DNA Methylation to ensure Faithful Epigenetic Memory Inheritance. *Genetics & epigenetics*, **2009**, 29-36.
18. Hashimoto, H., Horton, J.R., Zhang, X. and Cheng, X. (2009) UHRF1, a modular multi-domain protein, regulates replication-coupled crosstalk between DNA methylation and histone modifications. *Epigenetics*, **4**, 8-14.
19. Hashimoto, H., Vertino, P.M. and Cheng, X. (2010) Molecular coupling of DNA methylation and histone methylation. *Epigenomics*, **2**, 657-669.
20. Ward, D.C., Reich, E. and Stryer, L. (1969) Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *The Journal of biological chemistry*, **244**, 1228-1237.
21. Fiebig, T., Wan, C. and Zewail, A.H. (2002) Femtosecond charge transfer dynamics of a modified DNA base: 2-aminopurine in complexes with nucleotides. *Chemphyschem*, **3**, 781-788.

22. Jean, J.M. and Hall, K.B. (2001) 2-Aminopurine fluorescence quenching and lifetimes: role of base stacking. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 37-41.
23. Jean, J.M. and Hall, K.B. (2002) 2-Aminopurine electronic structure and fluorescence properties in DNA. *Biochemistry*, **41**, 13152-13161.
24. Jean, J.M. and Hall, K.B. (2004) Stacking-unstacking dynamics of oligodeoxynucleotide trimers. *Biochemistry*, **43**, 10277-10284.
25. O'Neill, M.A. and Barton, J.K. (2002) 2-Aminopurine: a probe of structural dynamics and charge transfer in DNA and DNA:RNA hybrids. *Journal of the American Chemical Society*, **124**, 13053-13066.
26. Wan, C., Fiebig, T., Schiemann, O., Barton, J.K. and Zewail, A.H. (2000) Femtosecond direct observation of charge transfer between bases in DNA. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 14052-14055.
27. Delagoutte, B., Lallous, N., Birck, C., Oudet, P. and Samama, J.P. (2008) Expression, purification, crystallization and preliminary crystallographic study of the SRA domain of the human UHRF1 protein. *Acta crystallographica*, **64**, 922-925.
28. Vuilleumier, C., Maechling-Strasser, C., Gerard, D. and Mely, Y. (1997) Evidence and prevention of HIV-1 nucleocapsid protein adsorption onto fluorescence quartz cells. *Anal Biochem*, **244**, 183-185.
29. Lakowicz, J.R. (1999) *Principles of fluorescence spectroscopy, Second edition*. Academic/Plenum Publishers, New York.
30. Lipari, G. and Szabo, A. (1980) Effect of librational motion on fluorescence depolarization and nuclear magnetic resonance relaxation in macromolecules and membranes. *Biophysical journal*, **30**, 489-506.
31. Kenfack, C.A., Piemont, E., Ben Gaied, N., Burger, A. and Mely, Y. (2008) Time-resolved fluorescent properties of 8-vinyl-deoxyadenosine and 2-amino-deoxyribosylpurine exhibit different sensitivity to their opposite base in duplexes. *The journal of physical chemistry. B*, **112**, 9736-9745.
32. Guest, C.R., Hochstrasser, R.A., Sowers, L.C. and Millar, D.P. (1991) Dynamics of mismatched base pairs in DNA. *Biochemistry*, **30**, 3271-3279.
33. Avilov, S.V., Godet, J., Piemont, E. and Mely, Y. (2009) Site-specific characterization of HIV-1 nucleocapsid protein binding to oligonucleotides with two binding sites. *Biochemistry*, **48**, 2422-2430.
34. Ramreddy, T., Rao, B.J. and Krishnamoorthy, G. (2007) Site-specific dynamics of strands in ss- and dsDNA as revealed by time-domain fluorescence of 2-aminopurine. *The journal of physical chemistry. B*, **111**, 5757-5766.
35. Avilov, S.V., Piemont, E., Shvadchak, V., de Rocquigny, H. and Mely, Y. (2008) Probing dynamics of HIV-1 nucleocapsid protein/target hexanucleotide complexes by 2-aminopurine. *Nucleic acids research*, **36**, 885-896.
36. Bloom, L.B., Otto, M.R., Beechem, J.M. and Goodman, M.F. (1993) Influence of 5'-nearest neighbors on the insertion kinetics of the fluorescent nucleotide analog 2-aminopurine by Klenow fragment. *Biochemistry*, **32**, 11247-11258.
37. Darii, M.V., Cherepanova, N.A., Subach, O.M., Kirsanova, O.V., Rasko, T., Slaska-Kiss, K., Kiss, A., Deville-Bonne, D., Reboud-Ravaux, M. and Gromova, E.S. (2009) Mutational analysis of the CG recognizing DNA methyltransferase SssI: insight into enzyme-DNA interactions. *Biochimica et biophysica acta*, **1794**, 1654-1662.
38. Holz, B., Klimasauskas, S., Serva, S. and Weinhold, E. (1998) 2-Aminopurine as a fluorescent probe for DNA base flipping by methyltransferases. *Nucleic acids research*, **26**, 1076-1083.

39. Neely, R.K., Daujotyte, D., Grazulis, S., Magennis, S.W., Dryden, D.T., Klimasauskas, S. and Jones, A.C. (2005) Time-resolved fluorescence of 2-aminopurine as a probe of base flipping in M.HhaI-DNA complexes. *Nucleic acids research*, **33**, 6953-6960.
40. Allan, B.W. and Reich, N.O. (1996) Targeted base stacking disruption by the EcoRI DNA methyltransferase. *Biochemistry*, **35**, 14757-14762.
41. Bheemanaik, S., Bujnicki, J.M., Nagaraja, V. and Rao, D.N. (2006) Functional analysis of amino acid residues at the dimerisation interface of KpnI DNA methyltransferase. *Biological chemistry*, **387**, 515-523.
42. Liebert, K., Hermann, A., Schlickerrieder, M. and Jeltsch, A. (2004) Stopped-flow and mutational analysis of base flipping by the Escherichia coli Dam DNA-(adenine-N6)-methyltransferase. *Journal of molecular biology*, **341**, 443-454.
43. Malygin, E.G., Evdokimov, A.A., Zinoviev, V.V., Ovechkina, L.G., Lindstrom, W.M., Reich, N.O., Schlagman, S.L. and Hattman, S. (2001) A dual role for substrate S-adenosyl-L-methionine in the methylation reaction with bacteriophage T4 Dam DNA-[N6-adenine]-methyltransferase. *Nucleic acids research*, **29**, 2361-2369.
44. Neely, R.K., Tamulaitis, G., Chen, K., Kubala, M., Siksnyis, V. and Jones, A.C. (2009) Time-resolved fluorescence studies of nucleotide flipping by restriction enzymes. *Nucleic acids research*, **37**, 6859-6870.
45. Reddy, Y.V. and Rao, D.N. (2000) Binding of EcoP15I DNA methyltransferase to DNA reveals a large structural distortion within the recognition sequence. *Journal of molecular biology*, **298**, 597-610.
46. Su, T.J., Connolly, B.A., Darlington, C., Mallin, R. and Dryden, D.T. (2004) Unusual 2-aminopurine fluorescence from a complex of DNA and the EcoKI methyltransferase. *Nucleic acids research*, **32**, 2223-2230.
47. Szegedi, S.S., Reich, N.O. and Gumpert, R.I. (2000) Substrate binding in vitro and kinetics of RsrI [N6-adenine] DNA methyltransferase. *Nucleic acids research*, **28**, 3962-3971.
48. Bonnist, E.Y., Liebert, K., Dryden, D.T., Jeltsch, A. and Jones, A.C. (2012) Using the fluorescence decay of 2-aminopurine to investigate conformational change in the recognition sequence of the EcoRV DNA-(adenine-N6)-methyltransferase on enzyme binding. *Biophysical chemistry*, **160**, 28-34.
49. Lenz, T., Bonnist, E.Y., Pljevaljcic, G., Neely, R.K., Dryden, D.T., Scheidig, A.J., Jones, A.C. and Weinhold, E. (2007) 2-Aminopurine flipped into the active site of the adenine-specific DNA methyltransferase M.TaqI: crystal structures and time-resolved fluorescence. *Journal of the American Chemical Society*, **129**, 6240-6248.
50. Bianchi, C. and Zangi, R. (2012) How to Distinguish methyl-Cytosine from Cytosine with High Fidelity. *Journal of molecular biology*.

## Supplementary information

### Materials and methods

#### Materials

The sequences of the 2-Ap-substituted oligonucleotides (ODNs) where the guanine facing the 2-Ap residue were replaced by a thymine within the sequence were 5'-GGGC2<sub>Ap</sub>XGCAGGG-3'/5'-CCCTGYGTGCC-3' for the 5T-duplexes, 5'-GGGCC2<sub>Ap</sub>GCAGGG-3'/5'-CCCTGCTGGCCC-3' for the NM6T-duplexes, and 5'-GGGCCXG2<sub>Ap</sub>AGGG-3'/5'-CCCTTYGGGCC-3' for the 8T-duplexes. The single CpG site was either unmethylated (NM: X = Y = C), hemi-methylated (HM: X = 5mC, Y = C), or fully methylated (BM: X = Y = 5mC).

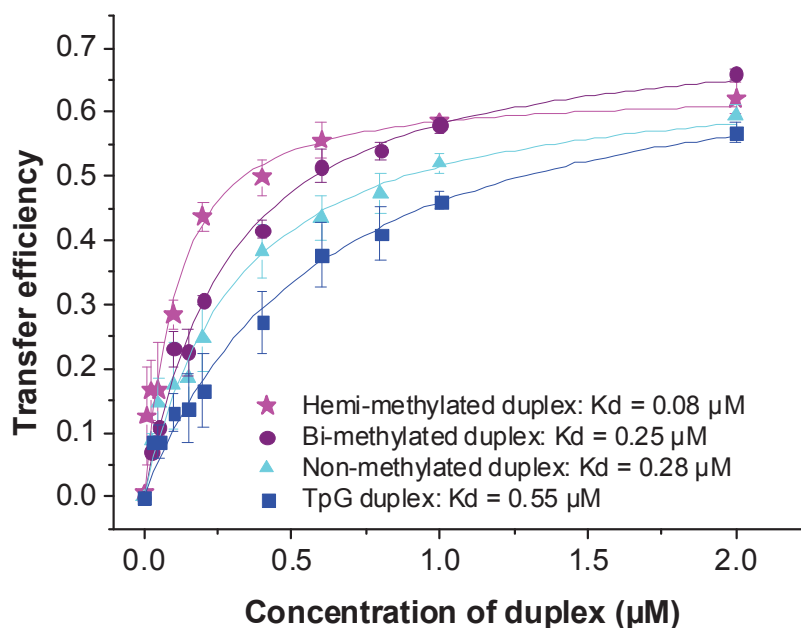
#### Fluorescence Resonance Energy Transfer (FRET)

The SRA (408-643) domain was labeled with Cy3 (Amersham Cy3 Maleimide Monoreactive dye 5-pack) on Cys 497 and purified by gel filtration (Superdex 200). ODNs were labeled with Cy5 at the 5' end (IBA GmbH Nucleic Acids Product Supply). Experiments were performed in 25 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.5 mM PMSF, 2 mM TCEP, 0.04 % PEG 20.000 at 20°C. Increasing concentrations of Cy5-labeled duplexes from 0 to 2 μM were added to a fixed concentration of 0.5 μM of Cy3-labeled SRA domain. Excitation wavelength is 500 nm and emission wavelengths are 520 to 800 nm. Transfer efficiency is measured from the relative fluorescence intensity of the donor (Cy3) in the absence ( $F_D$ ) and the presence of the acceptor (Cy5) ( $F_{DA}$ ) by:  $E = 1 - F_{DA}/F_D$  and plotted as a function of the duplex concentration. The theoretical curve is adjusted to the experimental points by the equation adapted from Didier et al (1):

$$E = E_f \left\{ \frac{1 + K_a ([SRA]_{tot} + [ADN]_{tot}) - \left\{ [1 + K_a ([SRA]_{tot} + [ADN]_{tot})]^2 - 4K_a^2 [SRA]_{tot} [ADN]_{tot} \right\}^{1/2}}{2K_a [SRA]_{tot}} \right\}$$

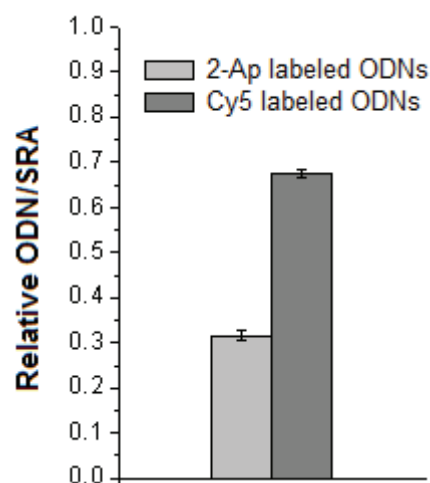
where  $[SRA]_{tot}$  and  $[ADN]_{tot}$  are the total concentrations of SRA protein and duplex, respectively,  $E_F$ , is the final value of FRET efficiency, and  $K_a$ , is the affinity constant of the protein.

## Results



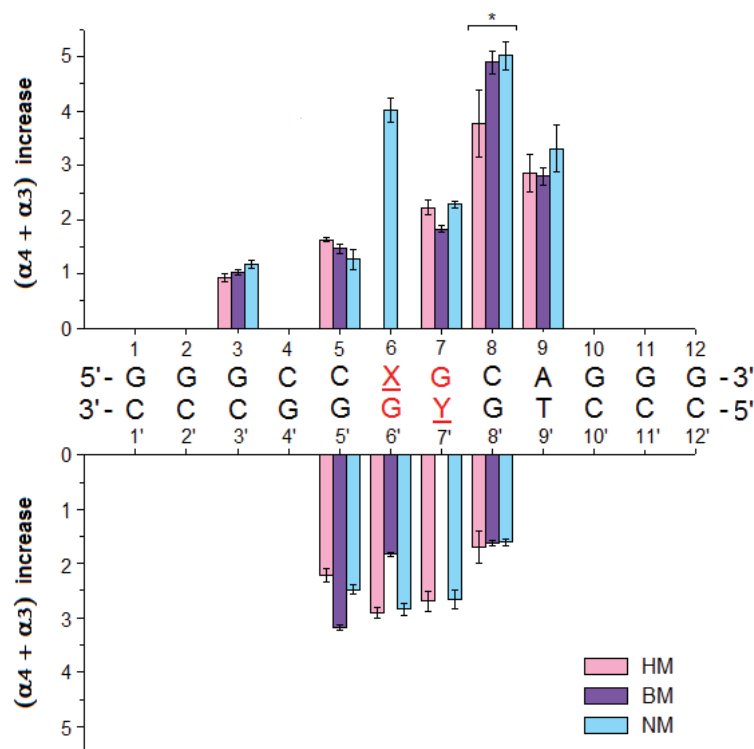
**Figure S1.** Titration of the SRA domain by DNA duplexes, using FRET in solution. Experiments were performed in 25 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.5 mM PMSF, 2 mM TCEP, 0.04 % PEG 20,000 at 20°C. Increasing concentrations of Cy5-labeled duplexes were added to a fixed concentration of 0.5 µM of Cy3-labeled SRA domain. Mean values and standard deviations for at least three experiments are reported on the graph. Except for the TpG duplex (squares) where a thymine substitutes the cytosine in the CpG site of the reference sequence, the duplexes used in this experiment contain a single CpG site that is either non-methylated (triangles), hemi-methylated (stars), or fully-methylated (circles). The  $K_d$  values of the SRA protein for each of these duplexes are indicated on the graph.





**Figure S2.** Competition between 2-Ap-labeled duplex ODNs and Cy5-labeled duplex ODNs for binding to the SRA domain. Relative ODN/SRA ratios are shown for Cy5-labeled and 2-Ap-labeled duplex ODNs in direct competition. These ratios were determined from the transfer efficiency between 0.5  $\mu$ M of Cy3-labeled SRA domain and 0.5  $\mu$ M of Cy5-labeled duplex ODNs in the presence of 0.5  $\mu$ M of 2-Ap labeled duplex ODNs. Experiments were performed three times in 25 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.5 mM PMSF, 2 mM TCEP, 0.04 % PEG 20,000 at 20°C. Mean values and standard deviations are reported on the histograms. The SRA domain binds preferentially to the Cy5-labeled duplexes, indicating the substitution of a natural base by a 2-Ap residue slightly decreases the SRA binding affinity.





**Figure S3.** Changes of  $(\alpha_3 + \alpha_4)$  lifetime amplitudes of the 2-Ap-labeled duplex ODNs on addition of the SRA (408-643) domain. The 12-bp duplex contains a single CpG recognition site (in red) that is either unmethylated (NM: X = Y = C), hemi-methylated (HM: X = 5mC, Y = C), or fully methylated (BM: X = Y = 5mC). The 2-Ap residue substitutes the natural base at various positions (3, 5, 6, 7, 8, 9, 5', 6', 7', 8') within the ODN sequence. Experiments were performed with 3  $\mu$ M of 2-Ap-labeled ODNs in the absence or the presence of 4  $\mu$ M SRA domain in 25 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.5 mM PMSF, 2 mM TCEP, 0.04 % PEG 20,000 at 20°C. Excitation and emission wavelengths were 315 nm and 370 nm, respectively. The mean values and standard deviations were calculated from the data of at least two experiments.

\* The results obtained with a sequence where the G facing the 2-Ap residue was replaced by a T revealed an increased sensitivity of 2-Ap at position 8 to the binding of SRA that may be related to the fact that 2-Ap does not base pair efficiently with G8 (2), so that a larger population can be driven to an extra-helical conformation as a consequence of SRA binding.

**Table S1.** Steady-state and time-resolved fluorescence parameters of 2-Ap-substituted non-methylated duplexes

	Q	$\alpha_0$	$\tau_1$ (ns)	$\alpha_1$	$\tau_2$ (ns)	$\alpha_2$	$\tau_3$ (ns)	$\alpha_3$	$\tau_4$ (ns)	$\alpha_4$	$\langle \tau \rangle$ (ns)
Free 2-Ap	0.680 <sup>a</sup>								10.2	1.00	10.2
NM3	0.067	0.28	0.05	0.53	0.66	0.06	3.35	0.05	9.01	0.08	1.39
NM3 + SRA	0.069	0.25	0.06	0.52	0.63	0.06	3.21	0.09	9.00	0.08	1.39
NM5	0.032	0.46	0.13	0.39	0.65	0.08	3.12	0.05	9.34	0.03	0.97
NM5 + SRA	0.037	0.39	0.08	0.42	0.55	0.09	3.05	0.07	9.61	0.03	0.91
NM6	0.012	0.70	0.07	0.20	0.46	0.06	3.26	0.02	9.51	0.01	0.73
NM6 + SRA	0.055	0.40	0.15	0.35	0.70	0.12	3.38	0.09	8.69	0.04	1.37
NM7	0.018	0.33	0.07	0.42	0.46	0.21	2.53	0.02	9.52	0.01	0.40
NM7 + SRA	0.028	0.35	0.09	0.40	0.50	0.18	2.82	0.06	9.52	0.01	0.64
NM8	0.010	0.72	0.05	0.22	0.55	0.03	3.41	0.02	9.20	0.01	0.62
NM8 + SRA	0.056	0.63	0.10	0.15	0.78	0.07	3.34	0.09	8.51	0.05	2.26
NM9	0.007	0.74	0.07	0.22	0.42	0.02	3.18	0.01	10.2	0.01	0.43
NM9 + SRA	0.016	0.69	0.07	0.24	0.58	0.03	3.03	0.04	10.3	0.01	0.77
NM8'	0.018	0.24	0.05	0.65	0.45	0.07	3.10	0.04	8.76	0.01	0.37
NM8' + SRA	0.028	0.19	0.09	0.64	0.51	0.10	2.96	0.06	9.48	0.02	0.52
NM7'	0.017	0.74	0.09	0.18	0.61	0.03	3.70	0.03	9.07	0.01	1.09
NM7' + SRA	0.049	0.62	0.14	0.17	0.93	0.09	3.40	0.09	8.86	0.04	1.92
NM6'	0.006	0.76	0.05	0.21	0.47	0.02	2.98	0.01	9.88	0.00	0.38
NM6' + SRA	0.012	0.74	0.07	0.19	0.58	0.03	3.15	0.04	10.4	0.00	0.75
NM5'	0.011	0.73	0.05	0.21	0.41	0.03	3.17	0.01	9.17	0.01	0.67
NM5' + SRA	0.021	0.59	0.06	0.31	0.46	0.05	3.08	0.05	9.39	0.01	0.75

Experiments were performed as for hemi-methylated duplex ODNs in Table 1.

<sup>a</sup>Data from (3)

**Table S2.** Steady-state and time-resolved fluorescence parameters of 2-Ap-substituted fully-methylated duplexes

	Q	$\alpha_0$	$\tau_1$ (ns)	$\alpha_1$	$\tau_2$ (ns)	$\alpha_2$	$\tau_3$ (ns)	$\alpha_3$	$\tau_4$ (ns)	$\alpha_4$	$\langle \tau \rangle$ (ns)
Free 2-Ap	0.680 <sup>a</sup>								10.2	1.00	10.2
BM3	0.068	0.38	0.06	0.40	0.68	0.07	3.37	0.07	8.52	0.09	1.64
BM3 + SRA	0.063	0.32	0.06	0.45	0.63	0.07	3.13	0.09	8.35	0.07	1.40
BM5	0.039	0.48	0.11	0.35	0.42	0.10	2.81	0.03	9.29	0.05	1.15
BM5 + SRA	0.050	0.32	0.08	0.45	0.50	0.10	2.95	0.07	9.37	0.05	1.12
BM7	0.035	0.27	0.06	0.40	0.43	0.25	2.89	0.05	8.51	0.03	0.72
BM7 + SRA	0.051	0.40	0.11	0.31	0.64	0.15	3.12	0.11	8.75	0.03	1.27
BM8	0.011	0.55	0.04	0.38	0.42	0.04	3.06	0.02	9.19	0.01	0.40
BM8 + SRA	0.059	0.61	0.08	0.17	0.67	0.07	3.35	0.09	8.54	0.06	2.01
BM9	0.009	0.64	0.07	0.31	0.40	0.03	3.32	0.01	10.2	0.01	0.39
BM9 + SRA	0.017	0.63	0.07	0.28	0.46	0.04	3.00	0.04	10.1	0.01	0.73
BM8'	0.019	0.34	0.07	0.53	0.46	0.08	2.97	0.04	8.84	0.01	0.44
BM8' + SRA	0.032	0.23	0.11	0.60	0.63	0.09	3.04	0.07	9.15	0.02	0.63
BM6'	0.025	0.47	0.07	0.41	0.50	0.06	3.08	0.04	9.73	0.02	0.73
BM6' + SRA	0.042	0.48	0.09	0.32	0.58	0.09	3.06	0.08	9.39	0.03	1.21
BM5'	0.011	0.74	0.05	0.19	0.37	0.04	3.22	0.01	9.64	0.01	0.59
BM5' + SRA	0.025	0.65	0.08	0.22	0.64	0.05	3.15	0.05	9.22	0.02	1.07

Experiments were performed as for hemi-methylated duplex ODNs in Table 1.

<sup>a</sup>Data from (3)

**Table S3.** Fluorescence anisotropy decay parameters of 2-Ap-substituted non-methylated duplex oligonucleotides

	$\varphi_1$ (ns)	$\beta_1$	$\varphi_2$ (ns)	$\beta_2$	$\varphi_3$ (ns)	$\beta_3$	S	$\theta_0$ (degree)
Free 2-Ap	0.08	1.00					0	90.0
NM3	0.32	0.55	2.01	0.45			0.673	40.2
NM3 + SRA	0.25	0.39	2.50	0.35	16.7	0.26	0.782	32.1
NM5	0.35	0.70	2.27	0.30			0.549	48.6
NM5 + SRA	0.25	0.38	2.33	0.22	16.1	0.40	0.787	31.7
NM6	0.35	0.67	2.09	0.33			0.574	47.0
NM6 + SRA	0.39	0.19	3.60	0.17	16.4	0.64	0.901	21.1
NM7	0.69	0.61	1.95	0.39			0.625	43.5
NM7 + SRA	0.24	0.28	3.60	0.22	17.3	0.50	0.848	26.3
NM8	0.32	0.66	3.02	0.34			0.584	46.3
NM8 + SRA	0.27	0.17	2.31	0.23	17.4	0.60	0.911	19.9
NM9	0.23	0.91	3.11	0.09			0.291	65.8
NM9 + SRA	0.10	0.26	4.68	0.28	16.2	0.46	0.860	25.4
NM8'	0.36	0.61	2.40	0.39			0.623	43.6
NM8' + SRA	0.22	0.29	3.13	0.20	17.3	0.51	0.840	26.8
NM7'	0.28	0.60	1.84	0.40			0.631	43.2
NM7' + SRA	0.27	0.15	3.07	0.24	16.4	0.61	0.922	18.5
NM6'	0.27	0.83	2.22	0.17			0.417	57.2
NM6' + SRA	0.15	0.22	4.96	0.26	15.8	0.52	0.881	23.3
NM5'	0.15	0.79	2.04	0.21			0.462	54.3
NM5' + SRA	0.13	0.29	5.01	0.29	17.0	0.43	0.844	26.8

Experiments were performed as for hemi-methylated duplex ODNs in Table 2.

**Table S4.** Fluorescence anisotropy decay parameters of 2-Ap-substituted fully-methylated duplex oligonucleotides

	$\varphi_1$ (ns)	$\beta_1$	$\varphi_2$ (ns)	$\beta_2$	$\varphi_3$ (ns)	$\beta_3$	S	$\theta_0$ (degree)
Free 2-Ap	0.08	1.00					0	90.0
BM3	0.38	0.55	1.96	0.45			0.669	40.6
BM3 + SRA	0.39	0.36	1.96	0.35	15.9	0.29	0.802	30.5
BM5	0.43	0.64	2.07	0.36			0.597	45.5
BM5 + SRA	0.21	0.31	2.45	0.33	17.3	0.36	0.832	27.9
BM7	0.55	0.62	1.88	0.39			0.618	44.0
BM7 + SRA	0.23	0.20	4.97	0.16	17.5	0.65	0.897	21.6
BM8	0.19	0.62	1.91	0.38			0.613	44.4
BM8 + SRA	0.33	0.21	2.53	0.17	16.2	0.62	0.887	22.7
BM9	0.24	0.87	2.50	0.13			0.366	60.6
BM9 + SRA	0.14	0.31	3.85	0.20	16.7	0.49	0.829	28.2
BM8'	0.33	0.70	2.28	0.30			0.544	48.9
BM8' + SRA	0.18	0.22	1.48	0.23	16.5	0.55	0.883	23.1
BM6'	0.32	0.75	2.23	0.25			0.502	51.7
BM6' + SRA	0.22	0.33	2.45	0.16	17.0	0.51	0.821	28.8
BM5'	0.19	0.80	1.75	0.21			0.452	55.0
BM5' + SRA	0.12	0.26	4.73	0.25	17.1	0.49	0.857	25.6

Experiments were performed as for hemi-methylated duplex ODNs in Table 2.

**Table S5.** Steady-state and time-resolved fluorescence parameters of 2-Ap-substituted duplex oligonucleotides where the guanine facing the 2-Ap residue was replaced by a thymine

	Q	$\alpha_0$	$\tau_1$ (ns)	$\alpha_1$	$\tau_2$ (ns)	$\alpha_2$	$\tau_3$ (ns)	$\alpha_3$	$\tau_4$ (ns)	$\alpha_4$	$\langle \tau \rangle$ (ns)
Free 2-Ap	0.680								10.2	1.00	10.2
HM5T	0.017	0.34	0.05	0.56	0.46	0.05	2.80	0.03	8.96	0.01	0.41
HM5T + SRA	0.024	0.26	0.05	0.60	0.46	0.07	2.86	0.05	8.85	0.02	0.49
NM5T	0.017	0.25	0.05	0.65	0.48	0.05	2.82	0.03	9.00	0.01	0.33
NM5T + SRA	0.025	0.24	0.06	0.63	0.56	0.06	2.92	0.05	9.21	0.02	0.49
BM5T	0.016	0.22	0.05	0.69	0.42	0.05	2.77	0.02	8.92	0.01	0.30
BM5T + SRA	0.022	0.29	0.06	0.59	0.47	0.06	2.86	0.04	8.80	0.02	0.46
NM6T	0.003	0.84	0.04	0.13	0.35	0.02	2.60	0.01	8.80	0.00	0.28
NM6T + SRA	0.011	0.61	0.03	0.31	0.30	0.04	2.89	0.03	9.61	0.00	0.43
HM8T	0.008	0.52	0.02	0.43	0.33	0.03	2.80	0.02	8.46	0.01	0.25
HM8T + SRA	0.017	0.61	0.04	0.29	0.42	0.04	2.90	0.04	8.93	0.01	0.66
NM8T	0.007	0.70	0.02	0.25	0.35	0.03	2.89	0.02	8.00	0.01	0.37
NM8T + SRA	0.014	0.65	0.03	0.26	0.46	0.04	3.11	0.04	8.95	0.01	0.63
BM8T	0.008	0.60	0.02	0.35	0.31	0.03	2.83	0.01	8.81	0.01	0.31
BM8T + SRA	0.016	0.27	0.02	0.62	0.27	0.05	2.88	0.04	8.79	0.01	0.32

Experiments were performed as for hemi-methylated duplex ODNs in Table 1.

<sup>a</sup>Data from (3)

**Table S6.** Fluorescence anisotropy decay parameters of 2-Ap-substituted duplex oligonucleotides where the guanine facing the 2-Ap residue was replaced by a thymine

	$\varphi_1$ (ns)	$\beta_1$	$\varphi_2$ (ns)	$\beta_2$	$\varphi_3$ (ns)	$\beta_3$	S	$\theta_0$ (degree)
Free 2-Ap	0.08	1.00					0	90.0
HM5T	0.30	0.61	2.03	0.39			0.624	43.6
HM5T + SRA	0.22	0.31	2.23	0.23	17.4	0.47	0.834	27.8
NM5T	0.42	0.72	2.82	0.28			0.529	49.9
NM5T + SRA	0.29	0.23	1.15	0.22	17.4	0.55	0.880	23.4
BM5T	0.37	0.66	2.61	0.34			0.587	46.1
BM5T + SRA	0.41	0.24	3.10	0.16	16.8	0.60	0.873	24.1
NM6T	0.27	0.75	3.07	0.25			0.502	51.7
NM6T + SRA	0.12	0.21	7.26	0.20	17.1	0.58	0.886	22.8
HM8T	0.27	0.60	2.23	0.40			0.631	43.2
HM8T + SRA	0.13	0.22	3.23	0.20	17.6	0.58	0.883	23.1
NM8T	0.30	0.68	2.63	0.33			0.570	47.2
NM8T + SRA	0.15	0.29	2.32	0.14	16.7	0.58	0.844	26.8
BM8T	0.35	0.55	1.83	0.45			0.669	40.6
BM8T + SRA	0.21	0.23	5.04	0.23	17.1	0.54	0.877	23.7

Experiments were performed as for hemi-methylated duplex ODNs in Table 2.

1. Didier, P., Sharma, K.K. and Mély, Y. (2011) In Podjarny, A., Dejaegere, A. P. and Kieffer, B. (eds.), *Biophysical Approaches Determining Ligand Binding to Biomolecular Targets: Detection, Measurement and Modelling* The Royal Society of Chemistry, pp. 156-199.
2. Kenfack, C.A., Piemont, E., Ben Gaied, N., Burger, A. and Mely, Y. (2008) Time-resolved fluorescent properties of 8-vinyl-deoxyadenosine and 2-amino-deoxyribosylpurine exhibit different sensitivity to their opposite base in duplexes. *The journal of physical chemistry. B*, 112, 9736-9745.
3. Ward, D.C., Reich, E. and Stryer, L. (1969) Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *The Journal of biological chemistry*, 244, 1228-1237.

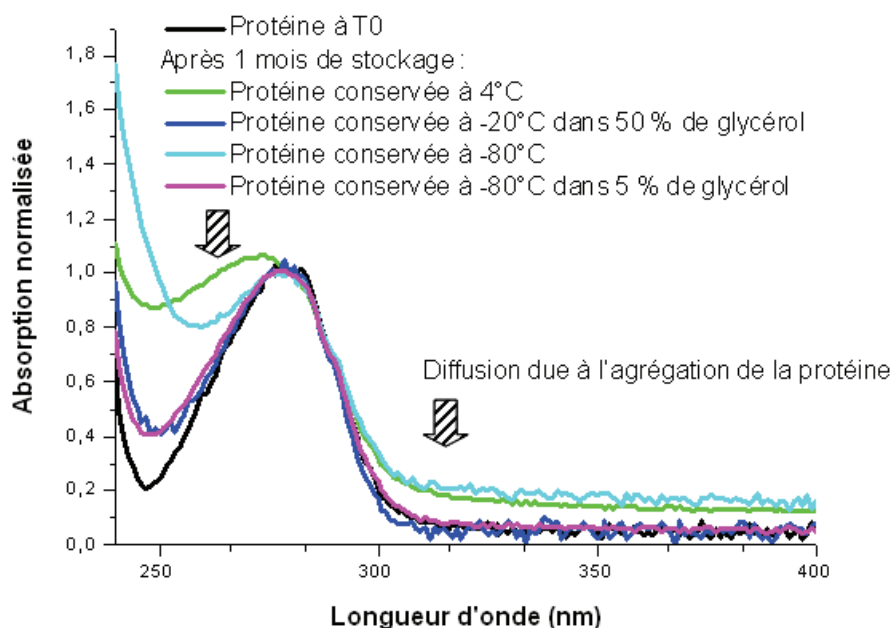


## B. Caractéristiques spectroscopiques du domaine SRA (408-643)

Les résidus aromatiques Trp sont des éléments fluorescents sensibles à l'environnement qui peuvent ainsi donner des informations sur la structure d'une protéine. Le domaine SRA (408-643) possédant cinq résidus Trp, la mesure de leurs spectres d'absorption, d'excitation et d'émission, ainsi que celles de leurs rendements quantiques et leurs temps de vie de fluorescence, nous ont permis de caractériser l'état conformationnel de la protéine et de confirmer le bon repliement du domaine SRA après production et purification de chaque nouveau lot.

### 1. Tests de conservation de la protéine

Lors du stockage du domaine SRA, nous avons pu observer une tendance de la protéine à précipiter. En effet, la mesure du spectre d'absorption de la protéine après un mois de stockage à 4°C ou -80°C révèle l'apparition d'une nette diffusion, témoin de l'agrégation des protéines (Figure 37). Afin d'éliminer ce phénomène, des tests de conservation ont été effectués en congelant la protéine dans du tampon contenant 5 ou 50 % de glycérol. Lors de la congélation, c'est généralement la formation de cristaux de glace qui endommage les protéines. L'addition de glycérol au tampon de conservation permet de diminuer la taille de ces cristaux et de limiter l'impact de la congélation sur la structure des protéines. Nos résultats révèlent que l'ajout de 50 ou 5 % de glycérol aux échantillons avant leur congélation à -20 ou -80°C, respectivement, permet de stabiliser la protéine (Figure 37). Une concentration trop importante de glycérol pouvant affecter les mesures de fluorescence, nous avons choisi de conserver le domaine SRA à -80°C dans du tampon contenant 5 % de glycérol. D'autre part, la congélation rapide des échantillons dans de l'azote liquide, ainsi que la préparation d'aliquots de petit volume évitant des cycles de congélation/décongélation, nous ont permis d'éviter l'agrégation de la protéine.



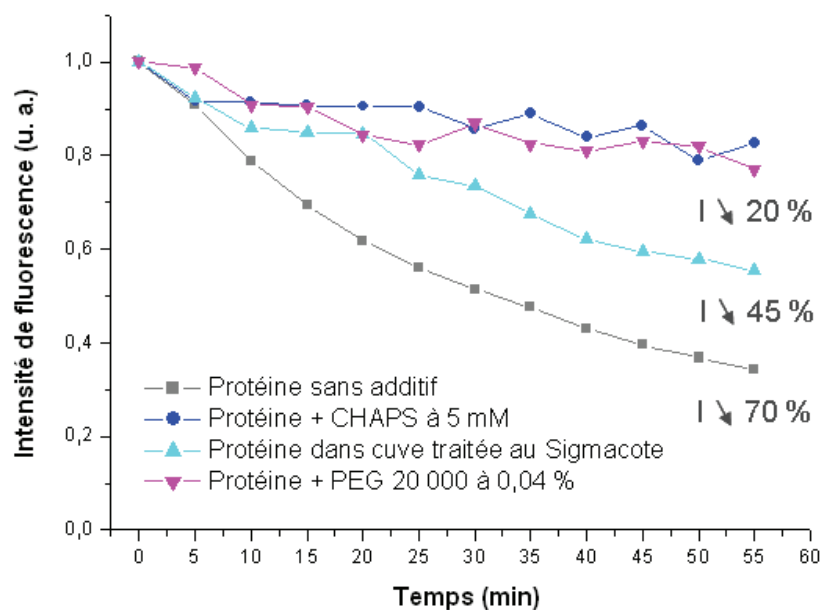
*Figure 37 : Spectre d'absorption du domaine SRA (408-643) lors des essais de conservation de la protéine. Les mesures sont effectuées sur 0,4  $\mu$ M de protéines à 20°C dans du tampon 25 mM Tris-HCl pH 7,5, 50 mM NaCl, 0,5 mM PMSF, 2 mM TCEP.*

## 2. Problèmes d'adsorption sur les parois des cuves de mesure

Lors des mesures de spectroscopie de fluorescence du domaine SRA, nous avons observé une diminution de l'intensité de fluorescence de la protéine au cours du temps, témoignant de l'adsorption progressive de la protéine sur les parois des cuves de mesure (Figure 38). Ainsi, en 55 minutes, l'intensité de fluorescence de la protéine est diminuée de près de 70 %. Afin de limiter ce phénomène, des mesures de l'intensité de fluorescence de la protéine ont été effectuées après ajout de divers agents anti-adsorbants au tampon de mesure. Le CHAPS (3-[(3-cholamidopropyl)diméthylammonio]-1-propanesulfonic acid) est un tensioactif zwitterionique permettant de diminuer la tension superficielle de surface des cuves (Ji et al., 2010). Le Sigmacote<sup>®</sup> (Sigma-Aldrich) est une solution de silicone permettant la formation d'un film protégeant la surface des cuves, et le PEG (Polyéthylène glycol) 20.000 est un polymère polyéther permettant également la formation d'un film protecteur sur les parois des cuves (Vuilleumier et al., 1997). Lorsque les cuves sont prétraitées avec du Sigmacote<sup>®</sup>, comparée à la baisse de 70 % qui était observée au bout de 55 minutes en absence d'additif, la diminution de l'intensité de fluorescence de la protéine n'est plus que de

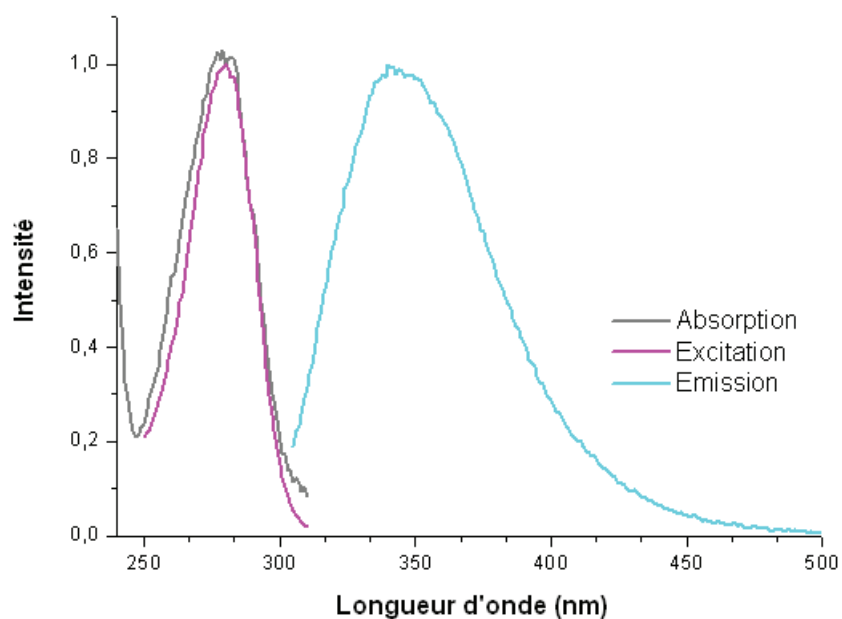
45 % (Figure 38). Ainsi, le prétraitement des cuves au Sigmacote® permet de diminuer la quantité de protéine adsorbée sur les parois des cuves. Cependant, ce processus d'absorption est encore plus largement diminué lorsque du PEG ou du CHAPS est ajouté au tampon de mesure, avec une baisse de seulement 20 % du maximum d'intensité de fluorescence de la protéine après 55 minutes d'incubation. Ces résultats indiquent que ces deux additifs sont de bons candidats pour limiter l'adsorption de la protéine sur les parois des cuves de mesure. Le PEG 20.000, peu coûteux, est finalement retenu comme additif anti-adsorbant.

L'ajout systématique de PEG 20.000 à 0,04 % dans le tampon de la protéine permet de diminuer significativement l'adsorption de la protéine sur les parois des cuves de mesure, mais ne l'élimine pas totalement. Par conséquent, l'intensité de fluorescence de la protéine est légèrement diminuée dans toutes les séries de mesures effectuées en spectroscopie de fluorescence.



*Figure 38 : Maximum d'intensité de fluorescence du domaine SRA (408-643) de la protéine UHRF1 au cours du temps en absence ou en présence de divers agents anti-adsorbants. Les mesures sont effectuées sur 0,4  $\mu$ M de protéines à 20°C dans du tampon 25 mM Tris-HCl pH 7,5, 50 mM NaCl, 0,5 mM PMSF, 2 mM TCEP. Le CHAPS est dilué à 5 mM dans le tampon pour protéine. Du PEG 20.000 à 4 % est pré-dissous dans de l'eau avant d'être ajouté à 0,04 % dans le tampon pour protéine. Pour le Sigmacote®, les cuves de mesures sont prétraitées selon le protocole indiqué par Sigma-Aldrich. La longueur d'onde d'excitation est de 295 nm.*

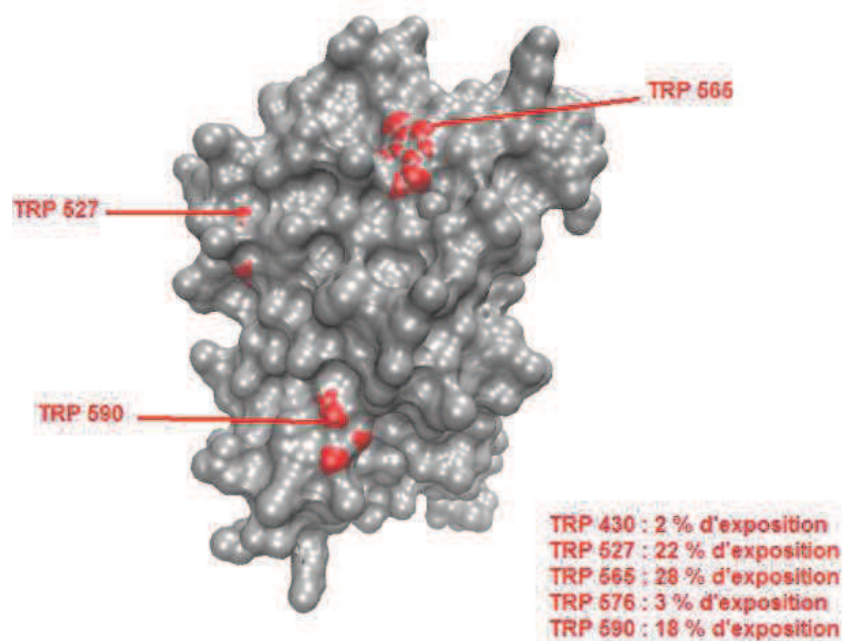
### 3. Caractéristiques spectroscopiques de la protéine



*Figure 39 : Spectres d'absorption, d'excitation et d'émission du domaine SRA (408-643) de la protéine UHRF1. Les mesures sont effectuées sur 0,4  $\mu$ M de protéines à 20°C dans du tampon 25 mM Tris-HCl pH 7,5, 50 mM NaCl, 0,5 mM PMSF, 2 mM TCEP, 0,04 % PEG 20.000. Les longueurs d'ondes d'excitation et d'émission sont de 295 nm et 320 nm respectivement.*

Afin de caractériser la protéine et de confirmer le bon repliement du domaine SRA, diverses mesures de spectroscopie de fluorescence ont été effectuées et réitérées après chaque cycle de production et de purification de la protéine. Le spectre d'absorption du domaine SRA présente une bande caractéristique de l'absorption du Trp et de Tyrosines (Figure 39). D'autre part, son spectre d'excitation enregistré à 320 nm se superpose au spectre d'absorption, indiquant l'absence de contaminant fluorescent dans l'échantillon. Le spectre de fluorescence de la protéine présente également une unique bande caractéristique de la fluorescence du Trp. La longueur d'onde du maximum d'émission de 344 nm révèle que les résidus Trp de la protéine se trouvent dans un environnement polaire. D'autre part, la valeur de 0,090 du rendement quantique de la fluorescence intrinsèque du domaine SRA (Tableau 2), largement inférieure à celle du Trp dans l'eau (0,14) (Eisinger and Navon, 1969), suggère que les résidus Trp de la protéine sont soit collectivement modérément fluorescents, soit que certains d'entre eux sont fortement éteints. La structure tridimensionnelle du domaine SRA (408-643) (Figure 40) révélant deux résidus Trp profondément enfouis dans la protéine, ce sont probablement

les trois résidus Trp exposés au solvant qui sont principalement responsables de la fluorescence observée de la protéine.



*Figure 40: Degrés d'exposition des résidus Trp de la protéine SRA (408-643). Seul trois résidus Trp sont visibles dans la représentation de surface du domaine SRA. Les deux autres Trp sont enfoui dans la protéine. Calculs effectués sur "<http://curie.utmb.edu/getarea.html>" et représentation sur VMD (code PDB 2PB7).*

*Tableau 2 : Paramètres de fluorescence du domaine SRA (408-643) de la protéine UHRF1.*

Q	$\tau_1$ (ns)	$\alpha_1$	$\tau_2$ (ns)	$\alpha_2$	$\tau_3$ (ns)	$\alpha_3$	$\langle \tau \rangle$ (ns)
0,090	0,19	0,27	0,97	0,29	3,11	0,44	1,71

*Les mesures sont effectuées sur 0,4  $\mu$ M de protéine à 20°C dans du tampon 25 mM Tris-HCl pH 7,5, 50 mM NaCl, 0,5 mM PMSF, 2 mM TCEP, 0,04 % PEG 20.000. Les longueurs d'ondes d'excitation et d'émission sont de 295 nm et 350 nm respectivement. Le meilleur ajustement aux déclin de fluorescence résolue en temps est obtenu avec trois temps de vie, où  $\tau_i$  correspond aux temps de vie,  $\alpha_i$  aux amplitudes relatives, et  $\langle \tau \rangle$  au temps de vie moyen. Q correspond au rendement quantique.*

Pour caractériser davantage la fluorescence intrinsèque du domaine SRA, des mesures de spectroscopie de fluorescence résolue en temps ont été effectuées sur la protéine. A une longueur d'onde d'émission de 350 nm, proche de la longueur d'onde du maximum d'émission de fluorescence de la protéine, le déclin de fluorescence résolue en temps est multi-exponentiel et comporte trois temps de déclin (Tableau 2). Le déclin de fluorescence est

dominé par le temps de vie long de 3,11 ns qui représente 44 % de la population totale. Les temps de déclin multiples des résidus Trp sont usuellement associés aux isomères conformationnels de ces résidus. L'observation de seulement trois composantes s'explique par les limites de la technique de déconvolution qui ne permet pas de séparer les temps de vie qui diffèrent de moins d'un facteur 1,5 (Bombarda et al., 1999; Mely et al., 1994). D'autre part, deux des résidus Trp sont profondément enfouis dans la protéine et ne sont probablement pas ou peu fluorescents (Figure 40).

D'une manière générale, les résultats obtenus en spectroscopie de fluorescence sont cohérents avec la structure tridimensionnelle du domaine SRA. Ces paramètres de fluorescence nous ont permis de vérifier l'intégrité de la protéine avant chaque série de mesures.

### C. Liaison du domaine SRA (408-643) à l'ADN

Bien que plusieurs études ont révélé une affinité préférentielle du domaine SRA pour l'ADN hémi-méthylé, les différences d'affinité de la protéine observées pour les divers types d'ADN, c'est à dire non-méthylés, bi-méthylés ou hémi-méthylés, sont variables d'une étude à l'autre (Arita et al., 2008; Avvakumov et al., 2008; Bostick et al., 2007; Frauer et al., 2011; Hashimoto et al., 2008; Qian et al., 2008; Rottach et al., 2010). Ainsi, afin de caractériser l'affinité du domaine SRA (408-643) avec l'ADN dans nos conditions, nous avons effectué des titrages par FRET en solution et par anisotropie de fluorescence.

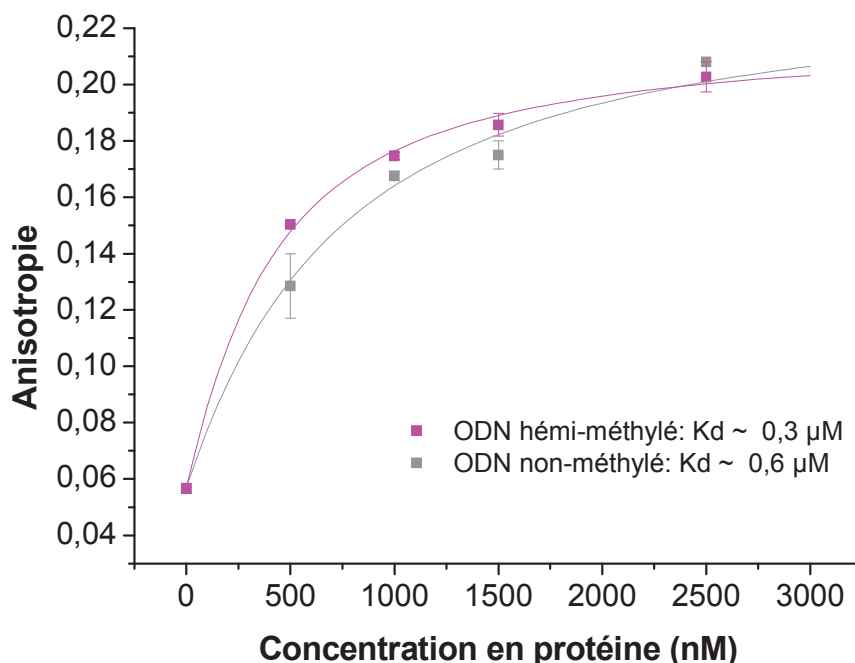
Les expériences de titrage par FRET en solution, en présence d'une concentration fixe de protéine marquée au Cy3 et de concentrations croissantes d'ODNs marqués au Cy5, ont confirmé l'affinité préférentielle du domaine SRA pour l'ADN hémi-méthylé (voir Figure S1 du Supplementary Data de la Publication 1). Le  $K_d$  de 80 nM obtenu pour l'ODN hémi-méthylé est du même ordre de grandeur que celui rapporté par Bostick et al. (Bostick et al., 2007). Cependant, alors que cette étude révèle une affinité du domaine SRA sept fois plus importante pour l'ADN hémi-méthylé que pour l'ADN bi-méthylé, nos résultats indiquent une préférence modérée de la protéine pour l'ADN hémi-méthylé d'un facteur quatre environ. D'autre part, nous avons pu observer des affinités similaires du domaine SRA pour l'ODN non-méthylé et l'ODN bi-méthylé, ainsi qu'une affinité non négligeable pour un ODN ne

comportant pas de site CpG. La faible préférence de la protéine pour l'ADN hémi-méthylé est en accord avec les études de Frauer et al. (Frauer et al., 2011), Qian et al. (Qian et al., 2008) et Rottach et al. (Rottach et al., 2010). En outre, une liaison du domaine SRA à un ADN ne comportant pas de site CpG avait été reportée dans l'étude de Frauer et al. (Frauer et al., 2011). Ces différences d'affinité peuvent en partie être expliquées par la variabilité des techniques (gel retard, anisotropie de fluorescence, test de liaison à l'ADN, FRET en solution) et du matériel (origine murine ou humaine du domaine SRA, taille du domaine SRA, taille des ODNs) utilisés dans ces différentes études.

En anisotropie de fluorescence, au contraire des expériences effectuées par FRET en solution, le titrage est en sens direct et une concentration croissante de protéine SRA est ajoutée à une concentration fixe d'ODNs (Figure 41). Dans ce contexte, on observe également une affinité préférentielle du domaine SRA pour l'ODN hémi-méthylé, et nos résultats sont en accord avec ceux obtenus par Qian et al. avec la même technique (Qian et al., 2008). Cependant, le  $K_d$  obtenu de  $0.3 \mu\text{M}$  est plus faible que celui révélé par les expériences de FRET en solution, et l'affinité de la protéine pour l'ODN hémi-méthylé est seulement deux fois plus importante que pour l'ODN non-méthylé. Ces différences d'affinité selon les techniques de mesure utilisées pourraient être expliquées par le mode de titrage direct ou indirect employé. En effet, dans les expériences de FRET en solution, une concentration fixe de protéine est titrée par une concentration croissante d'ODNs. Dans ce contexte, la concentration du domaine SRA est relativement faible pour les premiers points et la protéine se trouve donc en excès. L'adsorption de la protéine sur les parois des cuves est minimisée par la préparation séparée de chaque point de la courbe de titrage dans des tubes Eppendorf possédant une faible adhésion de surface ("low binding"). De plus, le domaine SRA est systématiquement ajouté à l'ODN, de manière à réduire l'adsorption de la protéine au maximum. Malgré ces précautions, la protéine se retrouve en excès pour les premiers points de la courbe de titrage, ce qui favorise son adsorption et entraîne une diminution de son intensité de fluorescence. Dans le cas des expériences d'anisotropie de fluorescence qui s'effectuent dans le cadre d'un titrage direct, ce sont les derniers points de la courbe qui sont limitant. En effet, le domaine SRA se retrouve très concentré pour les derniers points, et les mesures parallèles des spectres d'absorption révèlent l'apparition d'une diffusion, témoin de l'agrégation des complexes SRA/ODN formés. Il en résulte une augmentation artificielle de l'anisotropie dans les derniers points de titrage, le plateau n'étant jamais complètement atteint. Ainsi, les limitations des deux techniques expliquent la différence d'affinité observée par



FRET en solution et par anisotropie de fluorescence. Cependant, nos résultats montrent clairement que le domaine SRA possède une affinité préférentielle mais modérée pour l'ADN hémi-méthylé par rapport à l'ADN non-méthylé.



*Figure 41* : Mesure de l'interaction du domaine SRA (408-643) à l'ADN par anisotropie de fluorescence. Une concentration fixe de 0,2 µM d'oligonucléotide (ODN) est titrée par une concentration croissante de protéine. Les ODNs utilisés sont marqués à la fluorescéine et contiennent un site CpG hémi-méthylé ou non-méthylé. Les mesures sont effectuées à 20°C dans du tampon 25 mM Tris-HCl pH 7,5, 50 mM NaCl, 0,5 mM PMSF, 2 mM TCEP, 0,04 % PEG 20.000. La longueur d'onde d'excitation est de 470 nm. L'expérience est répétée au moins trois fois, les valeurs moyennes et les écarts types étant retranscrits sur le graphique. Les affinités de la protéine sont déterminées pour les divers types d'ODNs et les valeurs de  $K_d$  reportées sur le graphique.

Afin de déterminer la dépendance en sels de l'interaction entre le domaine SRA et l'ADN, nous avons effectué une série de titrages en anisotropie de fluorescence avec des tampons contenant diverses concentrations en NaCl (5, 50 ou 150 mM). L'affinité du domaine SRA pour les divers types d'ADN diminue faiblement en fonction de la concentration en sels. Malgré la charge positive du domaine SRA à pH neutre, son interaction avec l'ADN semble donc essentiellement de nature non-électrostatique.

Malgré la variabilité des affinités obtenues selon les techniques employées, la préférence de liaison du domaine SRA à l'ADN hémi-méthylé est donc bien établie. La préférence modérée de la protéine pour l'ADN hémi-méthylé révélée dans notre étude est en

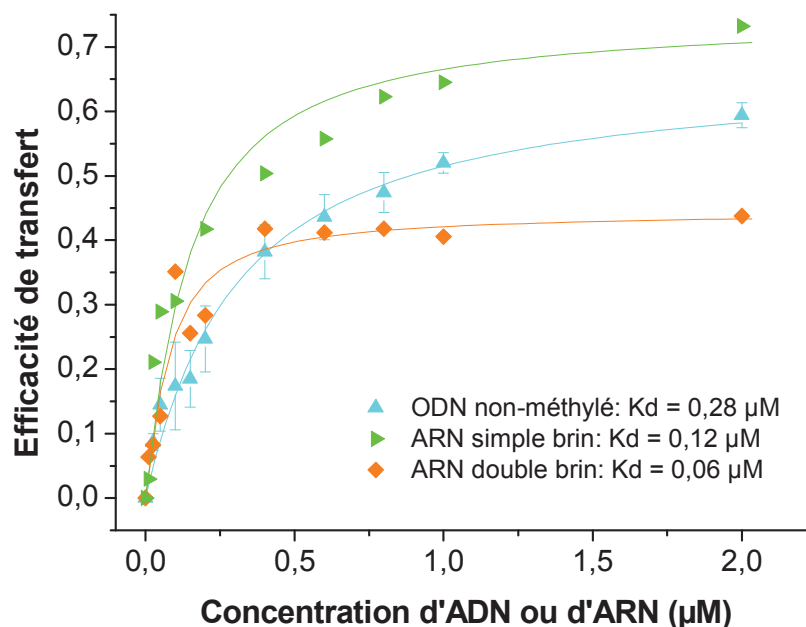
accord avec les observations de Frauer et al. (Frauer et al., 2011), Qian et al. (Qian et al., 2008) et Rottach et al. (Rottach et al., 2010). Ces résultats supportent l'hypothèse d'un rôle d'opérateur de lecture de la protéine UHRF1 qui serait capable de se lier à n'importe quel type d'ADN via son domaine SRA et de scanner les brins à la recherche de sites CpG hémi-méthylés.

#### D. Liaison du domaine SRA (408-643) à l'ARN

Plusieurs études récentes suggèrent que de petits ARNs pourraient servir de guide aux Dnmt3s afin de diriger la méthylation de l'ADN au niveau de sites spécifiques du génome (Aravin et al., 2008; Denis et al., 2011; Kuramochi-Miyagawa et al., 2008; Weinberg et al., 2006). Cependant, ce phénomène reste encore à éclaircir et l'interaction d'ARNnc avec la Dnmt1 n'a pas été observée. Afin d'établir si la protéine UHRF1, partenaire de la Dnmt1, pourrait jouer un rôle dans ces mécanismes de méthylation de l'ADN dirigée par l'ARN, nous avons effectué une série de titrages pour déterminer l'affinité du domaine SRA pour l'ARN simple et double brin. Ces expériences ont été réalisées en anisotropie de fluorescence et en FRET en solution mais dans des conditions non-optimales, c'est à dire sans limiter l'adsorption en préparant séparément chaque point de la courbe de titrage dans des tubes Eppendorf à faible adhésion de surface ("low binding"). En revanche, les résultats reportés dans la Figure 42 ont été obtenus par FRET en solution dans les bonnes conditions mais seulement sur une série de mesures.

Les titrages révèlent que le domaine SRA possède une forte affinité pour l'ARN simple et double brin, du même ordre que pour l'ADN hémi-méthylé. Cependant, en fonction du partenaire de la protéine, on observe des efficacités de transfert différentes dans les complexes formés à saturation (44 % pour l'ARN double brin, 73 % pour l'ARN simple brin, et ~ 62 % pour l'ADN). L'efficacité de transfert étant dépendante de la distance interchromophore, ces résultats indiquent que le domaine SRA se lierait aux ARNs dans des conformations ou des sites différents de ceux observés pour le complexe SRA/ADN. D'autre part, des expériences de compétition révèlent qu'une concentration d'ADN 20× supérieure à l'ARN ne suffit pas à déplacer la liaison formée entre l'ARN et le domaine SRA, l'inverse

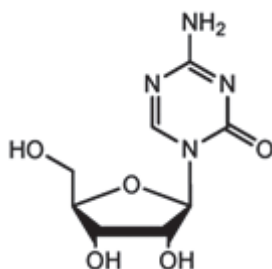
étant également observé. Ainsi, l'ARN et l'ADN semblent se lier sur deux sites différents du domaine SRA.



*Figure 42 : Mesure de l'interaction du domaine SRA (408-643) à l'ARN par FRET en solution. Une concentration fixe de  $0,05 \mu\text{M}$  de protéine marquée au Cy3 est titrée par une concentration croissante d'ARN simple ou double brin marqués au Cy5. Les ARNs utilisés possèdent une séquence homologue à l'ODN non-méthylé utilisé dans les expériences précédentes. Les mesures sont effectuées à  $20^\circ\text{C}$  dans du tampon  $25 \text{ mM Tris-HCl pH } 7,5$ ,  $50 \text{ mM NaCl}$ ,  $0,5 \text{ mM PMSF}$ ,  $2 \text{ mM TCEP}$ ,  $0,04 \%$  PEG 20.000. La longueur d'onde d'excitation est de  $500 \text{ nm}$ . L'efficacité de transfert est calculée selon la méthode décrite dans la section "Matériels et Méthodes". Les valeurs de  $K_d$  sont reportées sur le graphique.*

La forte affinité du domaine SRA observée pour l'ARN suggère que la protéine UHRF1 pourrait jouer un rôle encore inconnu dans la méthylation de l'ADN dirigée par l'ARN. En effet, la protéine UHRF1 étant impliqué dans le recrutement de la Dnmt1, la liaison simultanée de son domaine SRA à des ARNnc pourrait permettre de diriger l'enzyme au niveau de régions particulières du génome afin d'induire leur méthylation. Bien que cette piste soit intéressante et ouvre la porte sur une fonction encore inconnue de la protéine UHRF1, des expériences supplémentaires seront nécessaires afin d'éclaircir ces résultats préliminaires.

## E. Effet de la 5-azacytidine sur la liaison SRA/ADN

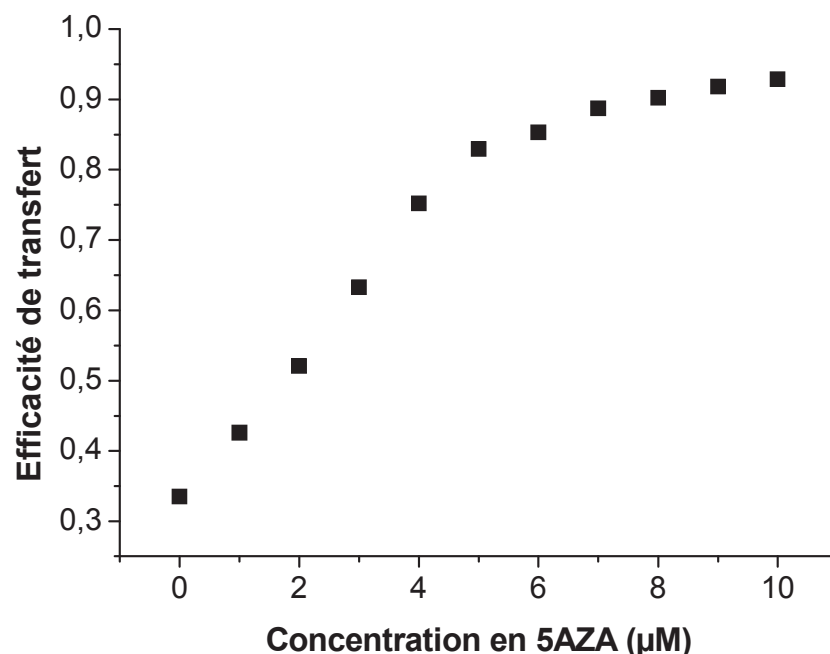


*Figure 43: Structure chimique de la 5-azacytidine (5-AZA).*

La 5-azacytidine (5-AZA) (Figure 43) est un inhibiteur des Dnmts utilisé dans les traitements anticancéreux. Analogue nucléosidique, cette base modifiée est incorporée à l'ADN à la place du résidu cytosine lors de la réplication et empêche la méthylation de l'ADN du fait de la présence d'un groupement nitrogène remplaçant le carbone 5 de la cytosine. Le mécanisme d'action communément accepté du 5-AZA implique la formation d'un complexe covalent irréversible entre la base et la Dnmt1, entraînant la dégradation de l'enzyme par le protéasome (Ghoshal et al., 2005). Cependant, ce mécanisme ne suffit pas à expliquer la déméthylation considérable de l'ADN induite par cet agent anticancéreux. La protéine UHRF1 jouant également un rôle important dans l'établissement de la méthylation de l'ADN par recrutement de la Dnmt1 au niveau des cytosines héli-méthylées reconnues par le domaine SRA, il serait possible que le 5-AZA puisse avoir un effet sur la liaison SRA/ADN. Afin de vérifier cette hypothèse, nous avons observé l'action du 5-AZA sur l'interaction SRA/ADN par des mesures de FRET en solution.

Après préformation d'un complexe SRA/ADN, du 5-AZA est ajouté en concentration croissante et l'efficacité de FRET est mesurée (Figure 44). Une nette augmentation de l'efficacité de transfert est observée lorsque le 5-AZA est ajouté en concentration croissante, jusqu'à l'obtention d'un plateau à 93 % d'efficacité de transfert. Ainsi, le 5-AZA ne déplace pas la liaison protéine/ADN comme un inhibiteur classique mais semble induire un changement de conformation du complexe entraînant une diminution de la distance interfluorophore. D'autre part, lorsque du 5-AZA est pré-incubé avec la protéine marquée au Cy3, la liaison d'ODNs marqués au Cy5 est toujours possible, l'efficacité de transfert observée dans ce contexte étant également augmentée par rapport au complexe formé en absence de 5-

AZA. Ces résultats révèlent sans ambiguïté l'existence d'une interaction entre le 5-AZA et le domaine SRA induisant des changements de conformation de la protéine mais n'excluant pas son interaction avec l'ADN. Etant donné la nature du 5-AZA qui est un analogue de la cytosine, il ne serait pas étonnant que cette base modifiée prenne la place de la cytosine dans la poche de liaison du domaine SRA en empêchant le basculement de base induit par la liaison de la protéine sur l'ADN. D'autre part, le 5-AZA pourrait bloquer la protéine, l'empêchant de glisser le long de l'ADN et de transférer les sites CpG hémi-méthylés à la Dnmt1. Le mécanisme d'action du 5-AZA au niveau de la liaison SRA/ADN reste à éclaircir mais il semble évident qu'il pourrait jouer un rôle dans l'inhibition de la méthylation de l'ADN observée après traitements au 5-AZA.



*Figure 44 : Efficacité de transfert de l'interaction protéine SRA/ADN en fonction de l'ajout en 5-AZA. 0,05 µM de protéine marquée au Cy3 est incubé pendant 2 minutes avec 0,2 µM d'ODN hémi-méthylé marqué au Cy5. Du 5-AZA est ajouté en concentration croissante au complexe SRA/ODN préformé. Les mesures sont effectuées à 20°C dans du tampon 25 mM Tris-HCl pH 7,5, 50 mM NaCl, 0,5 mM PMSF, 2 mM TCEP, 0,04 % PEG 20.000. La longueur d'onde d'excitation est de 500 nm. L'efficacité de transfert est calculée selon la méthode décrite dans la section "Matériels et Méthodes".*

## V- CONCLUSION ET PERSPECTIVES

La protéine UHRF1 se situe à l'interface des modifications de l'ADN et des histones, via son domaine SRA (Arita et al., 2008; Avvakumov et al., 2008; Bostick et al., 2007; Frauer et al., 2011; Hashimoto et al., 2008; Qian et al., 2008; Rottach et al., 2010; Sharif et al., 2007; Unoki et al., 2004) et ses domaines TTD et PHD (Arita et al., 2012; Nady et al., 2011; Xie et al., 2012). Ainsi, en permettant simultanément la lecture et l'écriture de modifications épigénétiques par recrutement de divers acteurs au niveau de régions spécifiques du génome, la protéine UHRF1 semble jouer un rôle central dans la maintenance et la transmission du code épigénétique (Bronner et al., 2010; Hashimoto et al., 2009; Hashimoto et al., 2010). L'étude de cette protéine apporte donc un nouveau niveau de compréhension dans la régulation des mécanismes épigénétiques. D'autre part, la protéine UHRF1 étant impliquée dans le développement de certaines tumeurs, elle pourrait constituer une cible thérapeutique alternative aux inhibiteurs des Dnmts utilisés dans les traitements anticancéreux.

La protéine UHRF1 joue un rôle important dans la transmission des profils de méthylation de l'ADN. En effet, lors de la réplication, en reconnaissant les sites CpG hémi-méthylés via son domaine SRA, la protéine UHRF1 permettrait de diriger la Dnmt1 au niveau de ces régions stratégiques afin de dupliquer les profils de méthylation sur le brin d'ADN nouvellement synthétisé (Bostick et al., 2007; Sharif et al., 2007). Dans ce contexte, les travaux effectués au cours de cette thèse ont permis de clarifier les mécanismes d'interaction du domaine SRA de UHRF1 avec l'ADN. Ainsi, l'étude des déclins de fluorescence de la 2-Ap, sensible à l'environnement, nous ont permis de caractériser la dynamique des acides nucléiques au sein des complexes SRA/ADN. Les résultats obtenus sur l'ADN hémi-méthylé en interaction avec le domaine SRA confirment la structure du complexe révélée par les études de cristallographie (Arita et al., 2008; Avvakumov et al., 2008; Hashimoto et al., 2008). Ainsi, le domaine SRA induirait le basculement de base de la cytosine méthylée tout en maintenant simultanément la structure globale de l'ADN. D'autre part, un ancrage particulier de la protéine au niveau du site CpG hémi-méthylé suggère que le domaine SRA induirait des changements de conformation pouvant favoriser la liaison de son partenaire Dnmt1 au niveau de la cytosine à méthyler sur le brin complémentaire. Alors que plusieurs études suggèrent un basculement de base spécifiquement induit par le domaine SRA sur la méthylcytosine de l'ADN hémi-méthylé (Arita et al., 2008; Avvakumov et al., 2008;

Hashimoto et al., 2008), nos travaux indiquent que la protéine adopterait le même type d'interaction avec l'ADN non-méthylé et bi-méthylé. Les similitudes de liaison du domaine SRA sur les différents types d'ADN, ainsi que la préférence modérée de la protéine pour l'ADN hémi-méthylé, supporte l'hypothèse d'un rôle d'opérateur de lecture de la protéine UHRF1. Ainsi, malgré sa préférence pour l'ADN hémi-méthylé, la protéine UHRF1 serait capable de se lier à n'importe quel type d'ADN via son domaine SRA et de scanner les brins à la recherche de sites CpG hémi-méthylés. Ce modèle reste à confirmer et des études d'interaction entre le domaine SRA et l'ADN par la technique de FRET en molécules uniques permettraient d'apporter des éléments de réponse afin de clarifier ce mécanisme. Les expériences de FRET en solution en présence de la protéine marquée au Cy3 et d'ODNs marqués au Cy5 ont démontré que le domaine SRA de UHRF1 interagissait bien avec l'ADN hémi-méthylé avec une efficacité de transfert significative. Grâce à cette technique, nous avons également pu observer une affinité du domaine SRA avec l'ADN non méthylé, l'ADN bi-méthylé, l'ADN ne comportant pas de site CpG, et l'ARN simple ou double brin, la protéine semblant adopter une conformation différente en fonction de l'identité de son partenaire, ADN ou ARN. Ces résultats suggèrent que l'interaction du domaine SRA de UHRF1 avec l'ADN ou l'ARN pourra être étudiée en molécule unique afin d'élucider la dynamique de la protéine dans ces différents complexes. Un projet de recherche a ainsi été conçu dans cette optique en collaboration avec l'équipe du Pr Alain Burger. Ce projet a reçu deux financements par la FRM et l'ANR intitulés respectivement "Epigénétique et méthylation de l'ADN: développement d'outils pour la compréhension du mécanisme de méthylation de l'ADN impliquant UHRF1 et DNMT1" (FRM édition 2011) et "Analyse du mécanisme de méthylation de l'ADN impliquant UHRF1 et DNMT1 par de nouveaux outils en fluorescence" (ANR édition 2012). Ainsi, l'utilisation en molécule unique d'ODNs marqués au Cy5 ou avec un analogue nucléosidique fluorescent optimisé dérivé de la 3-hydroxy-chromone, permettra de mieux caractériser l'interaction du domaine SRA avec l'ADN. Des premiers essais encourageants ont été effectués et grâce à l'accumulation des données obtenues avec cette technique, de précieuses informations sur les modifications structurales du domaine SRA en présence de ses divers partenaires pourront être obtenues. Elle permettra ainsi de déterminer l'association, la dissociation et le temps de résidence du domaine SRA sur les divers ODNs. D'autre part, pour mieux comprendre comment le domaine SRA reconnaît les sites hémi-méthylés, des expériences seront effectuées pour étudier le déplacement de la protéine le long de l'ADN. Des ODNs de tailles variées contenant un site hémi-méthylé



proche de leur extrémité et un fluorophore à proximité seront utilisés. Lorsque le domaine SRA marqué sera ajouté, la distribution des histogrammes de FRET ("high FRET" et "low FRET") permettra de détecter le mouvement de la protéine le long de l'ADN afin de déterminer si la protéine reconnaît directement le site hémi-méthylé (une population de "high FRET") ou si elle glisse le long de l'ADN à la recherche de ce site (Liu et al., 2008). La comparaison des résultats obtenus avec l'ADN hémi-méthylé avec ceux observés sur les ADN non-méthylé et bi-méthylé, devrait permettre une description complète des mécanismes de liaison du domaine SRA à l'ADN. Enfin, l'introduction de la Dnmt1 à notre modèle d'étude permettra de compléter la compréhension des mécanismes de méthylation de l'ADN, afin de déterminer si l'enzyme déplace effectivement l'interaction entre le domaine SRA et l'ADN (Arita et al., 2008). Ce projet permettra donc d'avancer considérablement dans la compréhension des mécanismes de méthylation de l'ADN.

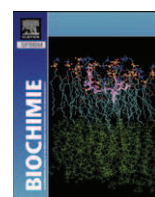
Les expériences de FRET en solution ont permis de mettre en évidence une interaction du 5-AZA, un analogue nucléosidique inhibiteur des Dnmts, avec le complexe SRA/ADN induisant un changement de conformation de la protéine. Une part de l'inhibition de la méthylation de l'ADN induite par le 5-AZA pourrait donc être la conséquence d'une action de la molécule sur UHRF1. Une étude plus poussée de la dynamique de cette interaction avec des ODNs marqués à la 2-Ap ou en molécule unique, permettra de compléter la compréhension du mécanisme d'action du 5-AZA. D'autre part, ces résultats indiquent que l'interaction SRA/ADN suivie par la technique de FRET pourrait constituer un test de screening permettant l'identification de nouveaux agents déméthylants de l'ADN potentiellement utilisables dans les thérapies anticancéreuses.

Lors des études d'affinité du domaine SRA, nous avons pu observer que la protéine présentait une affinité non négligeable pour l'ARN. Ces résultats préliminaires suggèrent qu'en se liant simultanément à la Dnmt1 et à l'ARN via son domaine SRA, la protéine UHRF1 pourrait jouer un rôle dans la méthylation de l'ADN dirigé par l'ARN. Les études de ces complexes en molécules uniques permettront certainement d'éclairer ce mécanisme. D'autre part, il serait intéressant de déterminer l'affinité du domaine SRA pour un triplex ADN:ARN. En effet, des études récentes ont révélés une affinité préférentielle de la Dnmt3b pour ce type de complexe comparé à l'ADN seul, qui permettrait de diriger la méthylation de l'ADN au niveau de régions spécifiques du génome (Schmitz et al., 2010). L'observation d'une préférence du domaine SRA pour un triplex ARN:ADN permettrait donc de supporter d'avantage l'hypothèse d'une implication de la protéine UHRF1 dans la méthylation de l'ADN

promue par l'ARN. D'autre part, des expériences d'immunoprécipitation à partir d'extraits cellulaires permettraient de confirmer l'interaction entre UHRF1 et l'ARN *in vitro*, et l'identification par RT-PCR des séquences d'ARN impliquées.

## VI- TRAVAUX ANNEXES

- A. Publication 2 : "Characterization of the lipid and protein organization in HBsAg viral particles by steady-state and time-resolved fluorescence spectroscopy"



## Research paper

# Characterization of the lipid and protein organization in HBsAg viral particles by steady-state and time-resolved fluorescence spectroscopy<sup>☆</sup>

Vanille J. Greiner<sup>a</sup>, Caroline Egelé<sup>a</sup>, Sule Oncul<sup>a</sup>, Frédéric Ronzon<sup>b</sup>, Catherine Manin<sup>b</sup>, Andrey Klymchenko<sup>a</sup>, Yves Mély<sup>a,\*</sup>

<sup>a</sup>Laboratoire de Biophotonique et Pharmacologie, UMR 7213 CNRS, Université de Strasbourg, Faculté de pharmacie, 74 route du Rhin, 67401 Illkirch, France

<sup>b</sup>Sanofi pasteur, 1541 avenue Marcel Mérieux, 69280 Marcy l'étoile, France

## ARTICLE INFO

## Article history:

Received 22 January 2010

Accepted 19 April 2010

Available online 24 April 2010

## Keywords:

HBsAg viral particles

Low density lipoprotein

Lipid membranes

Fluorescence spectroscopy

Fluorescent probes

## ABSTRACT

Hepatitis B surface antigen (HBsAg) particles, produced in the yeast *Hansenula polymorpha*, are 20 nm particles, composed of S surface viral proteins and host-derived lipids. Since the detailed structure of these particles is still missing, we further characterized them by fluorescence techniques. Fluorescence correlation spectroscopy indicated that the particles are mainly monomeric, with about 70 S proteins per particle. The S proteins were characterized through the intrinsic fluorescence of their thirteen Trp residues. Fluorescence quenching and time-resolved fluorescence experiments suggest the presence of both low emissive embedded Trp residues and more emissive Trp residues at the surface of the HBsAg particles. The low emission of the embedded Trp residues is consistent with their close proximity in alpha-helices. Furthermore, S proteins exhibit restricted movement, as expected from their tight association with lipids. The lipid organization of the particles was studied using viscosity-sensitive DPH-based probes and environment sensitive 3-hydroxyflavone probes, and compared to lipid vesicles and low density lipoproteins (LDLs), taken as models. Like LDLs, the HBsAg particles were found to be composed of an ordered rigid lipid interface, probably organized as a phospholipid monolayer, and a more hydrophobic and fluid inner core, likely composed of triglycerides and free fatty acids. However, the lipid core of HBsAg particles was substantially more polar than the LDL one, probably due to its larger content in proteins and its lower content in sterols. Based on our data, we propose a structural model for HBsAg particles where the S proteins deeply penetrate into the lipid core.

© 2010 Elsevier Masson SAS. All rights reserved.

## 1. Introduction

Hepatitis B virus (HBV) infection is one of the major causes of both acute and chronic liver hepatitis, liver cirrhosis and primary hepatocellular carcinomas [1,2]. Three types of particles are released from HBV infected cells: tubular lipoproteic structures with a 20–22 nm diameter, 22-nm spherical lipoproteic particles, and 40–44 nm infectious spherical virions, known as Dane particles

[3–7]. The 22-nm particles, called HB surface antigen (HBsAg) particles, correspond to the lipoproteic envelop of the mature infectious Dane particles. Moreover, the HBsAg particles found in the serum of patients with HBV infection outnumber mature virus particles, and are strongly immunogenic [8,9] so that they have been used to develop an effective vaccine against HBV [10]. Similar recombinant HBsAg particles are produced in different hosts, such as chinese hamster ovary cells [11], mouse fibroblast cells [12,13] or yeast cells [14], and have been shown to be immunogenically nearly identical to those purified from human serum [7].

HBsAg particles are composed of host-derived lipids and virus-encoded glycoproteins [12], with a protein to lipid weight ratio of about 60:40 [14]. The HBsAg glycoprotein exists in three forms designated L (large), M (medium) and S (small). The S protein is the major form, accounting for 80–90% of the total proteins [7,15]. Whereas HBsAg particles derived from human plasma contain two S proteins, p-25, the major protein, and gp-30, the glycosylated form of p-25 [16]; the *Hansenula polymorpha*-derived HBsAg particles used in this study contains only the nonglycosylated form

**Abbreviations:** AFM, atomic force microscopy; ApoB100, apolipoprotein B-100; DOPC, dioleoylphosphatidylcholine; DPH, diphenylhexatriene; EM, electron microscopy; ESIP, excited state-intramolecular proton transfer; FCS, fluorescence correlation spectroscopy; FRET, fluorescence resonance energy transfer; HBV, Hepatitis B virus; HBsAg, Hepatitis B surface antigen; LDL, low density lipoprotein; NATA, N-acetyl-L-tryptophanamide; PC, phosphatidylcholine; TMR, carboxy-tetramethylrhodamine; Trp, tryptophane.

<sup>☆</sup> This work was supported by a grant from Sanofi Pasteur.

\* Corresponding author. Tel.: +33 3 90 24 42 63; fax: +33 3 90 24 43 13.

E-mail address: [yves.mely@pharma.u-strasbg.fr](mailto:yves.mely@pharma.u-strasbg.fr) (Y. Mély).

of the S protein (24-kDa). About fifty dimers of protein S are found in the mice-derived HBsAg particles [13]. Alpha-helix is the most abundant secondary structure [17], accounting for 50% of the structure of HBsAg S proteins [7]. These proteins are highly hydrophobic, with only two relatively hydrophilic regions [17]. A model for the HBsAg protein organization has been proposed whereby regions within both the NH<sub>2</sub>- and COOH-terminal domains of p-25 and gp-30 are buried, while the antigenic 122–150 region is protruding out of the HBsAg particles, being thus exposed to the aqueous environment [16].

Together with the proteins, the lipids are also thought to be involved in the antigenic activity of HBsAg particles [7], likely by stabilizing the proper helical structure of the S proteins and the conformation of their hydrophilic region which contains the antigenic site [14]. The human and yeast HBsAg particles have the same composition including mainly phospholipids, with phosphatidylcholine (PC) being the main phospholipid [14,16], together with sterols, sterol esters and triacylglycerols. The majority of PC is accessible at the surface of the HBsAg particles, where they remain in a highly immobilized state in tight association with S proteins [12]. In addition, cryo-EM reconstitutions lead to the conclusion that unlike in mature virus particles, the HBsAg particles may contain the lipids in an unusual arrangement, being closely intercalated with the proteins [13]. Thus, it has been proposed that lipids in HBsAg particles are not aligned in a typical lipid bilayer configuration [12,14].

A powerful tool for characterization of lipid-based structures is fluorescent probes. For instance, information about the lipid order in the hydrophobic and the interfacial regions of lipid-based structures can be obtained respectively, by diphenylhexatriene (DPH) and its charged derivative (TMA-DPH), which are well-established probes of membrane fluidity [18,19]. Moreover, Nile Red shows high specificity to apolar lipid regions [20,21] and strong sensitivity to environment, providing information on the lipid organization. In addition, recently introduced membrane probes based on 3-hydroxyflavone provide the possibility for multiparametric probing of membrane environment [22]. Due to excited state-intramolecular proton transfer (ESIPT) reaction, they exhibit in the excited state, both a normal (N\*) and a tautomer (T\*) form, which are differently sensitive to their environment. In membranes, decomposition of their fluorescence spectra into N\*, T\* and additional hydrated H–N\* bands allows simultaneous probing of polarity and hydration [23]. Of particular interest for the present studies are the highly hydrophobic FN4 [24] and the charged derivative F2N8 [23,25] bearing an anchor group for membrane interface, which should allow specific probing of the hydrophobic and interfacial regions of the HBsAg lipid part, respectively. In this study, we used the intrinsic fluorescence of the Trp residues of the S proteins as well as fluorescent membrane probes to characterize the protein and lipid organization of HBsAg particles produced in the yeast *H. polymorpha*.

## 2. Materials and methods

### 2.1. Materials

HBsAg particles were obtained from Sanofi Pasteur. HBsAg was produced in the recombinant yeast *H. polymorpha* and obtained in a highly purified form by successive steps of fermentation, extraction and purification. Low density lipoproteins (LDL), Dioleoylphosphatidylcholine (DOPC), Dipalmitoylphosphatidylcholine (DPPC) and Nile Red were from Sigma–Aldrich. DPH and TMA-DPH were from Molecular Probes. 3-Hydroxyflavone dyes F2N8, FN4 and PPZ8 were synthesized as described elsewhere [24–26].

### 2.2. General procedure

All experiments were performed in 8 mM phosphate buffer saline (PBS) (NaCl 150 mM, pH 7.4), at 20 °C. Concentration of protein for HBsAg particles was 0.4 μM, except for quantum yield measurements, where it was 2 μM. For characterization of the lipid organization, the lipid concentration of HBsAg, LDLs and lipid vesicles was 6 μM. TMA-DPH and DPH were used at a 1 μM concentration. Staining of the particles and vesicles was performed by addition of an aliquot of a stock solution of the DPH or 3-hydroxyflavone dyes in DMSO.

### 2.3. Steady-state spectroscopy measurements

Absorption spectra were recorded on a Cary 400 spectrophotometer. Extinction coefficient of 81,650 M<sup>-1</sup> cm<sup>-1</sup> was used to determine the concentrations of S proteins in HBsAg particles at 280 nm. Fluorescence emission spectra were recorded at 20 °C on a FluoroMax spectrofluorometer (Jobin-Yvon) equipped with a thermostated cell compartment. Excitation wavelength was set at 295 nm to excite selectively the Trp residues. Spectra were corrected for the wavelength dependence of the emission monochromator and photomultiplier. Quantum yields were determined by taking *N*-acetyl-L-tryptophanamide (NATA) in water as a reference [27]. Quenching by acrylamide and KI was carried out by adding aliquots from stock solutions to the samples. The fluorescence intensity changes corrected for the dilution were recorded at the maximum emission wavelength. The data were analyzed by using Stern–Volmer equations. In the case of KI, quenching of fluorescence was described by:

$$F_0/F = 1 + K_{SV}[Q] \quad (1)$$

In the case of quenching by acrylamide, we used a modified equation:

$$F_0/F = (1 + K_{SV}[Q])\exp(V[Q]) \quad (2)$$

In these equations,  $F_0$  and  $F$  are the fluorescence intensities in the absence and the presence of quencher, respectively.  $[Q]$  is the concentration of quencher,  $K_{SV}$  is the Stern–Volmer quenching constant, and  $V$  is the static quenching constant. With both quenchers, the bimolecular quenching constant  $k_q$  was obtained by  $k_q = K_{SV}/\langle\tau\rangle$  where  $\langle\tau\rangle$  is the mean fluorescence lifetime in the absence of quencher.

Steady-state fluorescence anisotropy measurements were performed with a T-format SLM 8000 spectrofluorometer at 20 °C. The emitted light was monitored through interferential filters (373 ± 4 nm) (Schott, Mainz). A home-built device ensured the automatic rotation of the excitation polarizer.

### 2.4. Time-resolved fluorescence measurements

Time-resolved fluorescence measurements were performed with the time-correlated, single-photon counting technique, as previously described [28,29]. Excitation wavelength for Trp residues was at 295 nm. For lifetime measurements, the emission was collected through a polarizer set at magic angle (54.7°) and an 8 nm band-pass monochromator (Jobin-Yvon H10) at 320 nm, 340 nm, 360 nm and 380 nm. For time-resolved anisotropy measurements, this polarizer was set at the vertical position.  $I_{\perp}(t)$  and  $I_{\parallel}(t)$  were recorded alternatively every 5 s, by using the vertical polarization of the excitation beam with and without the interposition of a quartz crystal that rotates the beam polarization by 90°. Time-resolved data analysis was performed by the maximum entropy method using the Pulse5 software [30]. For the analysis of the fluorescence

decay, a distribution of 200 equally spaced lifetime values on a logarithmic scale between 0.01 and 10 ns was used.

The anisotropy decay parameters at 340 nm were extracted from both  $I_{\perp}(t)$  and  $I_{\parallel}(t)$  using:

$$I_{\parallel}(t) = I(t)[1 + 2r(t)]/3$$

$$I_{\perp}(t) = I(t)[1 - r(t)]/3$$

where  $I_{\parallel}$  and  $I_{\perp}$  are the intensities collected at emission polarizations parallel and perpendicular, respectively, to the polarization axis of the excitation beam.

Thus, the anisotropy at any time  $t$  is given by:

$$r(t) = \frac{I_{\parallel}(t) - GI_{\perp}(t)}{I_{\parallel}(t) + 2GI_{\perp}(t)} = r_0 \sum_i \beta_i e^{-t/\theta_i}$$

where  $r_0$  is the fundamental anisotropy,  $\beta_i$  corresponds to the fractional amplitude, which decays with the correlation time  $\theta_i$ , and  $G$  is the geometry factor at the emission wavelength.

### 2.5. FCS setup and data analysis

FCS measurements were performed on a two-photon platform including an Olympus IX70 inverted microscope, as described previously [31,32]. Two-photon excitation at 850 nm was provided by a mode-locked Tsunami Ti:sapphire laser pumped by a Millennia V solid state laser (Spectra Physics, U.S.A.). The measurements were carried out in an eight-well Lab-Tek II coverglass system, using a 400- $\mu$ L volume per well. The focal spot was set about 20  $\mu$ m above the coverslip. The normalized autocorrelation function,  $G(\tau)$  was calculated online by an ALV-5000E correlator (ALV, Germany) from the fluorescence fluctuations,  $\delta F(t)$ , by  $G(\tau) = \langle \delta F(t)\delta F(t + \tau) \rangle / \langle F(t) \rangle^2$  where  $\langle F(t) \rangle$  is the mean fluorescence signal, and  $\tau$  is the lag time. Assuming that PPZ8-labelled particles diffuse freely in a Gaussian excitation volume, the correlation function,  $G(\tau)$ , calculated from the fluorescence fluctuations was fitted according to [33]:

$$G(\tau) = \frac{1}{N} \left(1 + \frac{\tau}{\tau_d}\right)^{-1} \left(1 + \frac{1}{S^2} \frac{\tau}{\tau_d}\right)^{-1/2} \quad (3)$$

where  $\tau_d$  is the diffusion time,  $N$  is the mean number of molecules within the sample volume, and  $S$  is the ratio between the axial and lateral radii of the sample volume. The excitation volume is about 0.34 fl and  $S$  is about 3–4. Typical data recording times were 10 min, using HBsAg particles with a 0.4  $\mu$ M concentration of S proteins. HBsAg particles were fluorescently labeled by 10 min incubation with 0.4  $\mu$ M of the hydrophobic probe PPZ8. Using carboxytetramethylrhodamine (TMR) in water as a reference ( $D_{\text{TMR}} = 421 \mu\text{m}^2 \text{s}^{-1}$ ) [34], the diffusion coefficient,  $D_{\text{exp}}$ , of the labeled HBsAg particles was calculated by:  $D_{\text{exp}} = D_{\text{TMR}} \times \tau_{\text{d(TMR)}}/\tau_{\text{d}}$  (particles) where  $\tau_{\text{d(TMR)}}$  and  $\tau_{\text{d}}$  (particles) are the measured correlation times for TMR and PPZ8-labelled particles, respectively. The hydrodynamic diameter,  $d$ , of the HBsAg particles was calculated with the Stokes–Einstein equation:  $d = 2k_{\text{b}}T/6\pi\eta D_{\text{exp}}$ , where  $k_{\text{b}}$  is the Boltzmann constant,  $T$  is the absolute temperature and  $\eta$  is the viscosity of the solution.

## 3. Results and discussion

### 3.1. Steady-state fluorescence spectroscopy of the S proteins in HBsAg particles

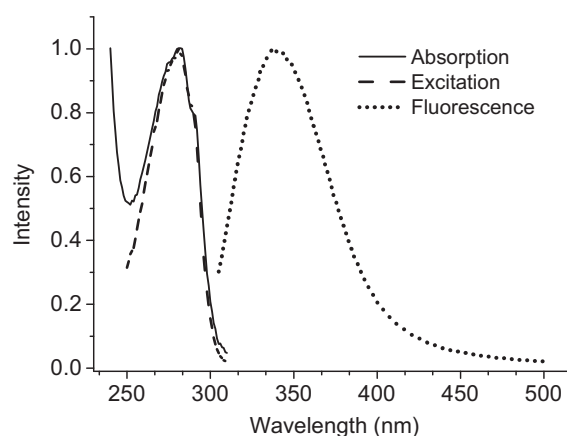
The absorption spectrum of HBsAg particles in solution (pH 7.4) shows one band characteristic for Trp and Tyr absorption (Fig. 1).

The fluorescence excitation spectrum recorded at 320 nm matches well to the absorption spectrum, showing that the Trp residues of the S proteins are the main emissive species in HBsAg particles. The fluorescence spectrum of HBsAg particles is also presented by a single band characteristic of Trp fluorescence (Fig. 1). The maximum emission wavelength (333 nm) was consistent with that reported in the literature [7,35] and indicated that on the average the Trp residues were in a rather low polar environment. Interestingly, the quantum yield of the intrinsic fluorescence of the HBsAg particles (0.058) was found to be less than half of that of free Trp in water (0.14) [27], suggesting that several Trp residues are likely weakly fluorescent.

To get further information, fluorescent quenching experiments were performed with the external KI quencher that selectively quenches the solvent-exposed Trp residues and the uncharged acrylamide quencher that can penetrate in the protein matrix and quench both solvent-exposed and buried Trp residues [36,37]. The Stern–Volmer quenching constants for KI and acrylamide were  $2.6 \pm 0.1 \text{ M}^{-1}$  and  $3.0 \pm 0.4 \text{ M}^{-1}$  (Fig. 2, A–B). From these values and the mean emission lifetime of the Trp residues (Table 1), the values of the apparent bimolecular quenching constants,  $k_{\text{q}}$ , were found to be of  $2.3 (\pm 0.1) \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$  for KI and  $2.7 (\pm 0.4) \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$  for acrylamide. These similar high  $k_{\text{q}}$  values for both quenchers are typical for solvent-exposed Trp residues [37,38] and indicate that most of the fluorescence comes from the Trp residues at the particle surface, exposed to solvent. Furthermore, the absence of significant deviation of the Stern–Volmer KI plot from linearity (Fig. 2A) further suggests that the low accessible embedded Trp residues are poorly fluorescent. Finally, since these exposed Trp residues show relatively low polar environment of their surrounding, we could speculate that these residues are close to the lipid–water interface of the HBsAg particles, thus ensuring low polarity, but high accessibility to the quenchers.

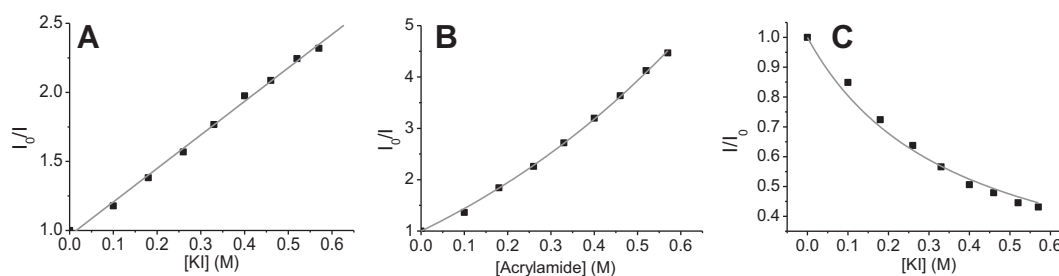
### 3.2. Time-resolved fluorescence of the S protein in particles

To further characterize the intrinsic fluorescence of the S proteins, the fluorescence decay parameters of the HBsAg particles were determined by the single-photon counting time-correlated technique at emission wavelengths of 320, 340, 360 and 380 nm. For all emission wavelengths, the time-resolved fluorescence decay of the S proteins is quite complex, showing four decay times. These multiple decay times are likely related to the large number of Trp



**Fig. 1.** Absorption, fluorescence excitation and emission spectra of HBsAg particles. For the emission spectrum, excitation wavelength was 295 nm; for the excitation spectrum, emission wavelength was 340 nm. Phosphate buffer 8 mM, NaCl 150 mM, pH 7.4 at 20 °C.





**Fig. 2.** Stern–Volmer quenching plots of the intrinsic fluorescence of HBsAg particles with KI and acrylamide. Solid lines correspond to the fit of the experimental points with KI (A) and acrylamide (B) to Equations (1) and (2), respectively, using the values of  $K_{SV}$  given in the text. For acrylamide, the static quenching constant  $V = 0.5 \pm 0.2 \text{ M}^{-1}$ . (C) Quenching data with KI fitted to the Stern–Volmer equation for a heterogeneous system (Equation (4)). Excitation and emission wavelengths are 295 and 340 nm, respectively.

residues in the S proteins. Indeed, individual Trp residues in proteins exhibit usually two to three lifetimes. As a consequence, with 13 Trp residues, as much as 26 to 39 lifetimes would be expected. However, the deconvolution technique does not allow to separate components that differ by less than a factor of 1.5 [28,39], explaining why the number of components is far less. Moreover, some of the Trp residues are probably not or poorly emissive.

At an emission of 340 nm, close to the maximum emission wavelength, the emission decay is dominated by the 0.85 ns component that represents 44%. Two additional components also strongly contribute to the decay: a shorter one (0.3 ns associated with an amplitude of 31%) and a longer one (2.2 ns associated with an amplitude of 20%). In contrast, the longer component (4.4 ns) is a minor component, representing only 5%.

The amplitude of  $\tau_1$  was highest at 320 nm, and strongly decreases at higher emission wavelengths, indicating that this lifetime corresponds mainly to embedded Trp residues in apolar environment. In contrast, the amplitudes associated with  $\tau_2$  and  $\tau_3$  were found to increase with the emission wavelength, suggesting that these lifetimes correspond to residues in more polar environment, and thus probably to residues more exposed to the solvent. The most exposed Trp population is likely associated with  $\tau_2$ , since its amplitude is highest at the more red shifted wavelength (380 nm) used in this study. Interestingly, the  $\tau_4$  lifetime value strongly increases with the emission wavelength, indicating that this lifetime probably describes the emission of both embedded and exposed Trp residues, and that the Trp residues associated with the longest-lived lifetimes correspond thus, to the most exposed ones.

Lifetime measurements allow refining KI quenching analyses. Indeed, knowing the lifetime values, the bimolecular quenching constants  $k_{qi}$  associated to each lifetime could be calculated, using [40]:

$$\frac{F}{F_0} = \sum_i \frac{f_i}{1 + k_{qi} \times \tau_i \times [Q]} \quad (4)$$

where  $f_i$  is the fractional contribution of the Trp residues associated to lifetime  $\tau_i$  to the steady-state intensity in the absence of quencher and  $[Q]$  is the concentration of quencher. According to

this equation, if we assume that the accessibility of the external KI quencher is negligible for the embedded Trp residues that mainly contribute to  $\tau_1$ , we found  $k_{q2} = 2.9 (\pm 0.3) \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ ,  $k_{q3} = 1.1 (\pm 0.1) \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$  and  $k_{q4} = 0.6 (\pm 0.1) \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$  (Fig. 2C). Of course, it should be kept in mind that the populations of Trp residues associated to each lifetime are heterogeneous, and thus, the  $k_{qi}$  values provide only a weighted average of the accessibility of the individual Trp residues contributing to each lifetime. Nevertheless, the  $k_{qi}$  values were in good agreement with the dependence of the lifetime amplitudes as a function of the emission wavelength, confirming that the most solvent-exposed Trp residues contribute strongly to the  $\tau_2$  lifetime and that the exposition degree of the lifetime-associated Trp populations follows the order:  $\tau_2 > \tau_3 > \tau_4 \gg \tau_1$ .

Since  $\tau_2$  and  $\tau_3$  lifetimes are the main contributors to the emission of the S proteins, this confirms that the more solvent-exposed Trp residues strongly contribute to the emission of the S proteins. According to the proposed model for the S proteins in membranes [41], these solvent-exposed Trp residues are likely Trp 156 and 223 on the hydrophilic face of the protein. In contrast, the short-lived lifetime  $\tau_1$  is mainly associated with embedded Trp residues. The poor fluorescence of the latter may be related to their close proximity in the hydrophobic alpha-helices that may result in efficient self-quenching [42,43]. For the same reason, the Trp pair (at positions 35 and 36) in the intra-helical region of the S proteins is likely poorly emissive, as well.

### 3.3. Steady-state and time-resolved fluorescence anisotropy of the S protein in particles

The steady-state anisotropy of the HBsAg particles was rather low ( $0.067 \pm 0.001$ ). At first glance, this low value was surprising since the S proteins are disulfide-bonded [44–46] and their motion is thought to be highly constrained in their lipidic environment [12]. In addition, the average fluorescence lifetime of the proteins is rather low (about 1 ns, see Table 1), which should further preclude a significant depolarization of the light. Accordingly, the anisotropy value suggests that the most emitting Trp residues may exhibit significant local and/or segmental motion. Moreover, the large

**Table 1**

Time-resolved fluorescence parameters of the HBsAg particles at several emission wavelengths. Excitation wavelength was set at 295 nm to excite selectively the Trp residues. The fluorescence lifetimes,  $\tau_i$ , the relative amplitudes,  $\alpha_i$ , the fractional intensities,  $f_i$ , and the mean lifetimes,  $\langle \tau \rangle$ , are expressed as means for at least three experiments. The standard deviations are usually below 15% for both the lifetimes and amplitudes, except for the longest-lived lifetime  $\tau_4$  which is ill defined.

$\lambda_{em}$ (nm)	$\tau_1$ (ns)	$\alpha_1$	$f_1$	$\tau_2$ (ns)	$\alpha_2$	$f_2$	$\tau_3$ (ns)	$\alpha_3$	$f_3$	$\tau_4$ (ns)	$\alpha_4$	$f_4$	$\langle \tau \rangle$ (ns)
320	0.30	0.45	0.16	0.81	0.37	0.34	1.90	0.13	0.28	3.70	0.05	0.21	0.87
340	0.31	0.31	0.09	0.85	0.44	0.33	2.20	0.20	0.39	4.40	0.05	0.19	1.13
360	0.28	0.24	0.06	0.80	0.47	0.31	2.30	0.25	0.47	5.00	0.04	0.16	1.22
380	0.23	0.18	0.03	0.80	0.57	0.36	2.70	0.25	0.54	7.70	0.01	0.06	1.26



number of Trp residues per protein probably allows homo-FRET between Trp residues, which also contributes to light depolarization [47]. To further characterize the system, time-resolved fluorescence anisotropy was performed.

Fluorescence anisotropy decays showed that S proteins were characterized by three correlation times (Table 2). The shortest correlation time was about 0.14 ns and can be assigned to the local motion of the Trp residues [36]. The 0.66 ns component likely corresponds to a segmental motion, which includes the most emitting Trp, and/or homo-FRET between Trp residues. The major component is clearly the 60 ns correlation time. However, this correlation time is an artifact since its value was found to follow the value of the correlation time limit that we provided to the software. Keeping in mind that correlation times that exceed 10 times the longest decay time could not be observed, we conclude that the real value of the long correlation time is larger than 60 ns. If we admit that S proteins are dimeric and thus have a molecular weight of 48 kDa [13,14], a correlation time of about 20 ns would have been expected for their overall rotation. The absence of this correlation time confirms that the S proteins exhibit a restricted motion in their lipid environment and that they may be assembled in high-order multimers in the particles [48]. To adequately model the anisotropy decay of the S proteins, we introduced thus, a limiting anisotropy,  $r_\infty$  that corresponds to the final anisotropy reached, after the molecules have emitted. We found an  $r_\infty$  value of 0.096 that is about half the 0.17 value obtained for the fundamental anisotropy  $r_0$ . In the case of Trp residues, a  $r_0$  value of 0.17 is usual at 295 nm excitation wavelength [36], due to the presence of two close excited states:  $^1L_a$  and  $^1L_b$ , which are characterized by different transition dipole moments. Comparison of  $r_\infty$  with  $r_0$  indicates that the motions of the Trp residues and homo-FRET decrease the anisotropy by about 40%. It can also be stressed that the introduction of  $r_\infty$  does not significantly modify the  $\theta_1$  and  $\theta_2$  values and their associated amplitudes. However, an additional component of 4 ns appeared that likely corresponds also to segmental motion and/or homo-FRET.

Taken together, our fluorescence data on the S proteins are consistent with the presence of embedded low emissive Trp residues together with more emissive Trp residues at the lipid–water interface of the HBsAg particles. The low emission of the embedded Trp residues may be a consequence of their close proximity, in line with inclusion of most of these Trp residues in alpha-helices. Furthermore, S proteins exhibit restricted movement, as expected from their tight association with lipids in the HBsAg particles.

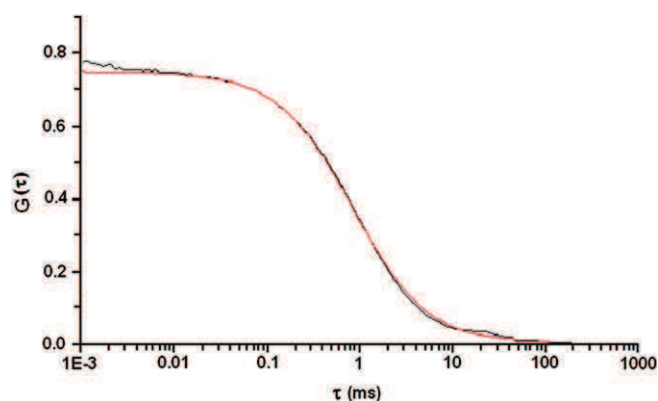
### 3.4. Fluorescence correlation spectroscopy

To get further insight in the organization of the HBsAg particles, we performed fluorescence correlation spectroscopy (FCS) using probe PPZ8, which showed strong affinity to HBsAg particles (see below). From the autocorrelation curve (Fig. 3), we obtained a diffusion constant of  $14.7 \pm 0.4 \mu\text{m}^2 \text{s}^{-1}$  for the labeled HBsAg particles, irrespective of the probe concentration. From this diffusion time, using the Stokes–Einstein equation assuming a spherical

**Table 2**

Time-resolved fluorescence anisotropy parameters of the S proteins in HBsAg particles. Excitation wavelength is 295 nm. The rotational correlation times,  $\theta_i$  and their relative amplitudes,  $\beta_i$  are expressed as means for at least three experiments. The standard deviations are usually below 20% for both parameters.  $r_0$  and  $r_\infty$  correspond to the fundamental and the limiting anisotropy, respectively.

$r_\infty$	$r_0$	$\theta_1$ (ns)	$\beta_1$	$\theta_2$ (ns)	$\beta_2$	$\theta_3$ (ns)	$\beta_3$
–	0.167	0.14	0.07	0.66	0.26	60.0	0.67
0.096	0.167	0.18	0.09	0.70	0.20	4.0	0.12



**Fig. 3.** PPZ8-labelled HBsAg particles, as monitored by fluorescence correlation spectroscopy. The autocorrelation curve was recorded with HBsAg particles (at a concentration of  $0.4 \mu\text{M}$ , as expressed in S proteins) labeled with  $0.4 \mu\text{M}$  PPZ8 in PBS buffer, pH 7.4, at  $20^\circ\text{C}$ . The black line is the average correlation curve and the red line is the fit to the experimental points with Equation (3). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

shape of the particles, we calculated a value of  $33 \pm 1 \text{ nm}$  for their hydrodynamic diameter, which is consistent with the 27 nm and 23 nm diameters of the particles measured by atomic force microscopy (AFM) and electron microscopy (EM), respectively [13,35].

From the mean measured number of particles in the focal volume provided by two-photon excitation and the concentration of S proteins measured by absorbance, the calculated number of proteins per particle is about 70, in good agreement with AFM data (manuscript in preparation), calculations from the cryo-EM data [13] or calculations based on the assumption that HBsAg particles have a mass of  $3.5 \times 10^6$  with 60% proteins (by mass).

### 3.5. Steady-state anisotropy with DPH and TMA-DPH

Information about the lipid order of the viral particles can be obtained by DPH and TMA-DPH, two well-established probes of membrane fluidity [18,19]. The orientation of DPH within lipid bilayers is relatively constrained and generally assumed to be parallel to the lipid acyl chain axis. TMA-DPH dye is a derivative of DPH, with a more defined localization in the membrane due to its cationic group that anchors the fluorophore at the lipid interface. Thus, TMA-DPH provides information on the particle interface more specifically than DPH.

To obtain information about their organization and lipid fluidity, HBsAg particles were characterized with both types of probes in comparison with model lipid vesicles, that present laterally organized lipid bilayers, and with LDLs that present a lipid monolayer encapsulating a core of cholesterol esters and triglycerides [48–50].

HBsAg particles exhibit a high anisotropy value for the surface probe TMA-DPH, close to that in gel phase vesicles and LDLs, suggesting a highly rigid interface. For the non charged DPH probe, the anisotropy value in HBsAg particles is identical to that in LDLs but much lower than in gel phase vesicles (Table 3). For both probes, the anisotropy values measured for HBsAg particles at  $20^\circ\text{C}$  are consistent with the values measured for the same particles at  $37^\circ\text{C}$  by Sonveaux et al. [48]; the slight differences being due to the expected decrease in fluidity when the temperature is lowered. The close anisotropy values shown by HBsAg particles and LDLs for both types of probes suggest a similar organization for the two types of particles, with a densely packed lipid interface and a more fluid inner core [52–54]. The similarities in LDL and HBsAg organization are in line with their close composition in phospholipids (with a majority of phosphatidylcholine) and triglycerides [14,52].

**Table 3**

Steady-state anisotropy of HBsAg particles, LDLs and lipid vesicle. Gel phase is represented by DPPC vesicles at room temperature, while fluid phase is represented by DOPC vesicles at 20 °C.

	Particles		Lipid bilayers	
	HBsAg	LDL	Gel	Fluid
TMA-DPH	0.303 (0.291) <sup>a</sup>	0.288 <sup>b</sup>	0.350	0.190
DPH	0.226 (0.204) <sup>a</sup>	0.225 <sup>b</sup>	0.330	0.110

<sup>a</sup> From ref. [48] at 37 °C.

<sup>b</sup> From ref. [51].

### 3.6. Characterization of the lipid organization of HBsAg particles with different fluorescent probes

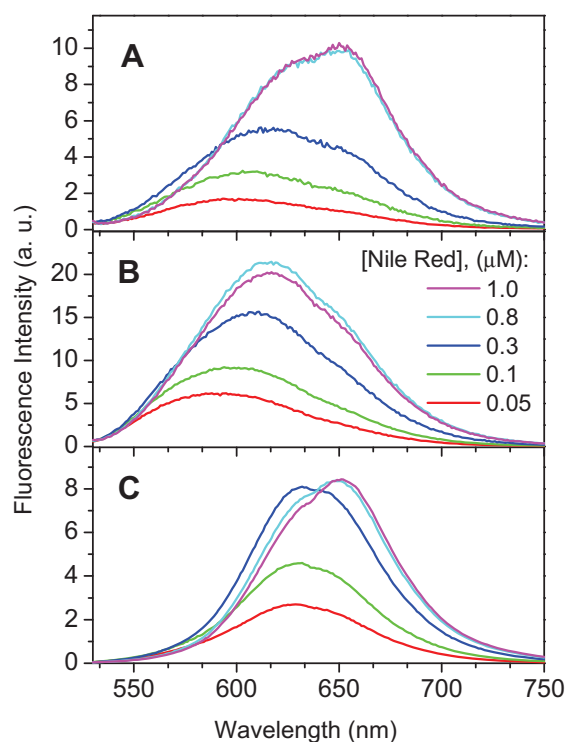
To further substantiate the similarities between HBsAg particles and LDLs, we then compared their lipid organization using different hydrophobic fluorescent probes sensitive to the environment: Nile Red and 3-hydroxyflavone derivatives. These probes bind to lipid-based structures and report on the local physicochemical properties of their binding site(s). Nile Red and FN4 are low polar probes that bind the most hydrophobic sites [20,21]. In contrast, dyes F2N8, bearing polar groups bind lipid structures at the lipid–water interface. Thus, utilization of both types of probes allows characterizing both the lipid membrane surface and more internal hydrophobic regions. Noticeably, the fluorescence of FN4 and F2N8 in water is almost negligible, while on binding to lipidic structures they become strongly fluorescent. Therefore, the observed fluorescence is mainly represented by the probes bound to the lipidic structures and is not significantly contaminated by the free probes in the buffer.

#### 3.6.1. Nile Red

The emission spectrum of Nile Red is highly sensitive to the polarity of its environment. This probe is low soluble and weakly fluorescent in water, but highly fluorescent being bound to lipidic structures [20,21]. We studied the fluorescence spectra of Nile Red at different concentrations in the presence of HBsAg particles, LDL particles and DOPC vesicles (Fig. 4).

In LDL and HBsAg particles, with increasing concentrations of Nile Red (0.05–0.3 μM) the fluorescence intensity increases gradually and the emission maximum progressively shifts to the red (Fig. 4). In contrast, for DOPC vesicles, the increase in the fluorescence intensity is not accompanied by a red shift in this concentration range. The absence of shift in DOPC vesicles is probably related to the homogeneity of the probe binding sites in the lipid bilayers. In contrast, LDLs and HBsAg particles show more heterogeneous binding sites for the probe. Indeed, the observed red shifts suggest that Nile Red molecules occupy first the most apolar regions of the particles and then, more polar regions. For LDLs, these observations are in agreement with their described structure, with namely an apolar core composed of cholesterol and triglycerides, surrounded by a more polar shell composed of a phospholipid monolayer containing unesterified cholesterol and the apolipoprotein B-100 [50]. The analogous behavior of Nile red in LDLs and HBsAg particles, confirm that the two particles exhibit comparable hydrophobic cores and lipid–water interfaces.

Additionally, the low fluorescence increase between 0.8 and 1 μM (Fig. 4) suggests that, for all batches, binding sites are saturated in this concentration range. This saturation behavior could be related to the high probe/lipid ratio (from 1/20 to 1/6) achieved in these conditions. It can be also noted that at high probe/lipid ratios, the Nile Red fluorescence in HBsAg particles and DOPC vesicles is close to that in water, indicating a large fraction of non-bound dyes in water. However, this is not the case for LDLs (Fig. 4), indicating



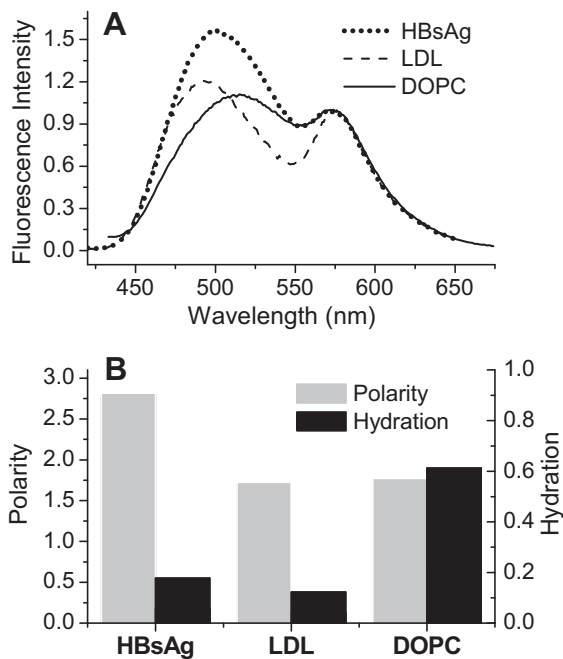
**Fig. 4.** Fluorescence emission spectra of Nile Red in HBsAg particles (A), LDL particles (B) and DOPC vesicles (C) at different concentrations of Nile red. The lipid concentration in LDLs, HBsAg particles and DOPC vesicles was ca 6 μM.

that these particles present a larger number of binding sites for Nile Red.

At low probe/lipid ratios, the maximum emission wavelength at 590 nm in LDLs is close to that reported for native LDLs (587 nm) [55]. In HBsAg particles, the emission maximum is slightly red shifted (600 nm), probably due to the larger protein content of HBsAg particles (60% in mass) [14] (data not shown), as compared to LDLs (20–25% in mass) [52,56], which contributes to the higher polarity of the probe binding sites. Importantly, in DOPC lipid bilayers, the emission maximum of Nile Red is strongly red shifted (627 nm) as compared to LDL and HBsAg particles. This confirms that in lipid vesicles, which do not contain a lipidic core, Nile Red binds mainly at the bilayer interface. In contrast, in LDL and HBsAg particles, Nile Red binds initially the low polar lipid core. At higher probe/lipid ratios, Nile Red binds more to the particle interface, and therefore the position of the emission maxima in LDL and HBsAg becomes closer to that in DOPC vesicles (Fig. 4).

#### 3.6.2. F2N8 probe

Due to its charged anchor group it binds preferentially to the lipid membrane interface [23], showing a well resolved dual emission in LDL, HBsAg particles, and DOPC vesicles (Fig. 5A). However all three spectra are significantly different from each other. Thus, the short-wavelength band for HBsAg and LDL particles is strongly blue shifted with respect to that for DOPC vesicles. Moreover, the relative intensity of the short-wavelength band for HBsAg particles is significantly higher than that for LDLs and DOPC vesicles. To understand better these spectroscopic differences, we deconvoluted the spectra into three bands: N\*, H–N\* and T\* (Fig. 5B). This analysis allows obtaining simultaneously parameters of polarity and hydration, which are two independent characteristics of the lipid assemblies [23,57]. The hydration parameter for LDLs is much lower than for DOPC vesicles, as expected from the

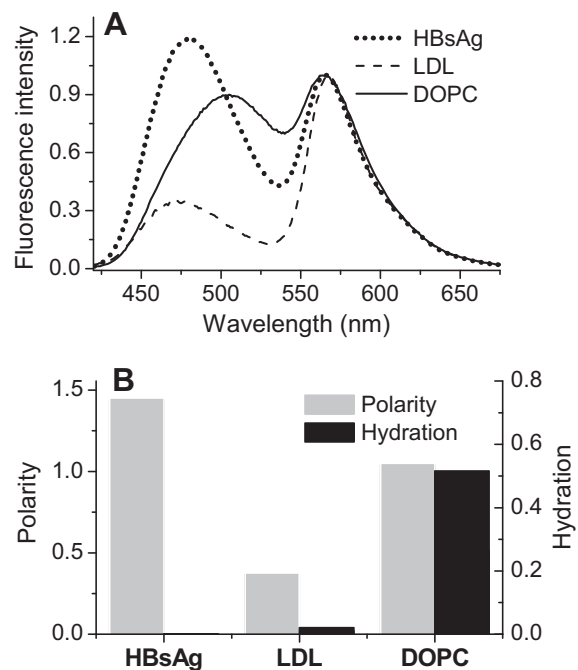


**Fig. 5.** Probing the polarity and hydration of F2N8 probe binding sites in HBsAg particles, LDL particles and DOPC vesicles. Fluorescence spectra (A). Polarity and hydration of F2N8 probe binding sites in the different lipid systems were estimated from deconvolution of its fluorescence spectra (B). Probe concentration was 0.05  $\mu\text{M}$ , while the lipid concentration for HBsAg particles, LDL particles and DOPC vesicles was 6  $\mu\text{M}$ .

tighter packing of the lipids at the LDL than at the DOPC surface, and the exclusion of water from the phospholipid monolayer by cholesterol [54,58]. Interestingly, the hydration parameter for HBsAg particles is comparable to that of LDLs, confirming that, like in LDLs, the lipids on the surface of the HBsAg particles are tightly packed and probably organized in a monolayer around the hydrophobic core. A bilayer structure at the HBsAg interface could be excluded, since the small size of the particles (20–30 nm) would impose a very strong curvature, and thus, a poor lipid packing density and strong hydration [23]. In contrast, the polarity parameter of HBsAg particles is much higher than in LDLs (Fig. 5B), likely as a consequence of the larger protein content of HBsAg particles as compared to LDLs.

### 3.6.3. FN4 probe

Similarly to Nile Red, the FN4 probe is highly hydrophobic and does not contain any charged group. Therefore, this dye should preferentially bind to highly hydrophobic sites. Probe FN4 also shows different spectroscopic profiles in LDLs, HBsAg particles and DOPC vesicles (Fig. 6A). Remarkably, the short-wavelength band for LDL and HBsAg particles is again significantly blue shifted with respect to that for DOPC vesicles. Moreover, the relative intensity of the short-wavelength band in HBsAg particles is much higher than in LDLs. Deconvolution analysis of these fluorescence spectra shows that LDLs and HBsAg particles are characterized by an almost absence of hydration, while in DOPC vesicles the hydration value is comparable to that observed with probe F2N8 (Fig. 6B). This confirms that the lipid organization in LDL and HBsAg particles largely differs from that in DOPC vesicles. In LDL and HBsAg particles, FN4 probably occupies their poorly hydrated core, explaining the absence of FN4 hydration. In contrast, in DOPC vesicles, which do not contain a lipid core, the probe bound to the lipid bilayer shows a hydration close to that for the interfacial probe F2N8. An additional remarkable observation is the dramatic



**Fig. 6.** Probing the polarity and hydration of FN4 probe binding sites in HBsAg particles, LDL particles and DOPC vesicles. Fluorescence spectra (A). Polarity and hydration of FN4 probe binding sites in the different lipid systems were estimated from deconvolution of their fluorescence spectra (B). Probe concentration was 0.05  $\mu\text{M}$ , while the lipid concentration for HBsAg particles, LDL particles and DOPC vesicles was 6  $\mu\text{M}$ .

difference in polarity between LDL and HBsAg particles. Indeed, the polarity parameter is nearly 4-times larger in HBsAg particles than in LDLs, while with F2N8 probe this difference was only 1.6 times. Taking into account that F2N8 binds at the interface, while FN4 binds at the most hydrophobic available sites, the differences in polarity between LDL and HBsAg particles appear much more pronounced in the hydrophobic core than at the interface. These differences are probably connected with a different organization of the cores. Indeed, in LDLs, the apolipoprotein B-100 is wrapped around the surface of the particle and thus, only partially penetrates into the inner core [50] while in HBsAg particles, the S proteins seem to protrude largely into the lipid core. Thus, the hydrophobic core of HBsAg particles contains a large number of S proteins while the LDL core is composed primarily of cholesterol esters [49]. Moreover, the larger sterol content of LDLs (40% in mass) [52] as compared to HBsAg particles (3% in mass) [14] likely also contributes to the lower polarity of the LDL core.

In conclusion, our fluorescence data allowed us to confirm and complete the data in the literature describing the structure of the HBsAg particles. In this respect, the value of the diameter of HBsAg particles that we determined by FCS (33 nm) is fully consistent with that obtained by AFM [35], EM [13,14,59] and dynamic light scattering [14]. Similarly, the number of proteins ( $\sim 70$ ) per particle deduced by FCS is in excellent agreement with that determined by cryo-EM [13] and AFM (data not shown). Furthermore, our data with the fluorescent lipidic probes suggest that the S proteins are largely immersed in the lipid core of the particle, in line with previous EM and low-angle X-ray scattering data [59]. The S proteins were found to exhibit a restricted motion, suggesting their tight association with the lipids in the HBsAg particles, in agreement with the electron spin resonance, EM and AFM observation that the lipids of the HBsAg particles are in a highly immobilized state [12,13,35]. Steady-state and time-resolved fluorescence spectroscopy data on the S proteins further revealed



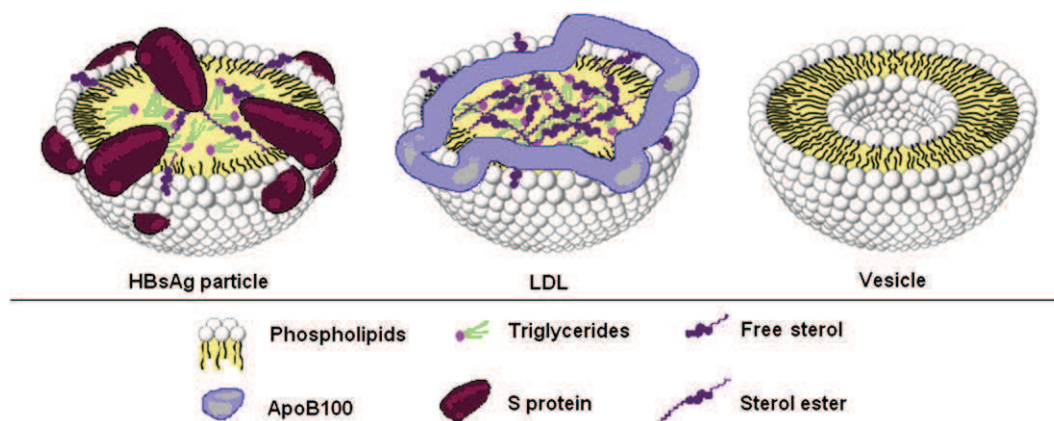


Fig. 7. Schematic model of lipid vesicle, LDL and HBsAg particle. The schematic model of LDL was adapted from Prassl et al. [50].

the presence of solvent-exposed Trp residues, that are likely associated with the antigenic region protruding out of the HBsAg particles [13,16,17,35,41]. Taken together, our data allow us to propose a model for the HBsAg structure in line with the schematic model of Satoh et al. [12] (Fig. 7). In this model, the HBsAg particles exhibit a lipoprotein-like structure in which a compact phospholipid monolayer surrounds a more hydrophobic and fluid inner core, likely composed of triglycerides, fatty acids and sterol esters. The major difference with LDLs is that in contrast to the apolipoprotein B-100 mainly located at the LDL surface [49,50], the S proteins are largely immersed in the lipid core, as suggested by the high polarity of the lipid core observed with the fluorescent probes. Moreover, HBsAg particles contain a larger quantity of protein (60% in mass) as compared to LDLs (20–25% in mass).

## Acknowledgements

The authors thank Guy Duportail and Pascal Didier for fruitful discussion.

## References

- [1] D. Ganem, H.E. Varmus, The molecular biology of the hepatitis B viruses. *Annu. Rev. Biochem.* 56 (1987) 651–693.
- [2] W.S. Robinson, L.I. Lutwick, The virus of hepatitis, type B (first of two parts). *N. Engl. J. Med.* 295 (1976) 1168–1175.
- [3] M.E. Bayer, B.S. Blumberg, B. Werner, Particles associated with Australia antigen in the sera of patients with leukaemia, Down's syndrome and hepatitis. *Nature* 218 (1968) 1057–1059.
- [4] A.I. Sutnick, W.T. London, B.S. Blumberg, Australia antigen and the quest for a hepatitis virus. *Am. J. Dig. Dis.* 14 (1969) 189–194.
- [5] R.J. Hirschman, N.R. Shulman, L.F. Barker, K.O. Smith, Virus-like particles in sera of patients with infectious and serum hepatitis. *JAMA* 208 (1969) 1667–1670.
- [6] D.S. Dane, C.H. Cameron, M. Briggs, Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. *Lancet* 1 (1970) 695–698.
- [7] F. Gaviñanes, J. Gomez-Gutierrez, M. Aracil, J.M. Gonzalez-Ros, J.A. Ferragut, E. Guerrero, D.L. Peterson, Hepatitis B surface antigen. Role of lipids in maintaining the structural and antigenic properties of protein components. *Biochem. J.* 265 (1990) 857–864.
- [8] W.J. McAleer, E.B. Buynak, R.Z. Maigetter, D.E. Wampler, W.J. Miller, M. R. Hilleman, Human hepatitis B vaccine from recombinant yeast. *Nature* 307 (1984) 178–180.
- [9] P. Maupas, A. Goudeau, P. Coursaget, J. Drucker, P. Bagros, Immunisation against hepatitis B in man. *Lancet* 1 (1976) 1367–1370.
- [10] J. Gomez-Gutierrez, I. Rodriguez-Crespo, D.L. Peterson, F. Gaviñanes, Antigenicity of hepatitis B surface antigen proteins reconstituted with phospholipids. *Biochim. Biophys. Acta* 1233 (1995) 205–212.
- [11] D. Diminsky, N. Moav, M. Gorecki, Y. Barenholz, Physical, chemical and immunological stability of CHO-derived hepatitis B surface antigen (HBsAg) particles. *Vaccine* 18 (1999) 3–17.
- [12] O. Satoh, H. Imai, T. Yoneyama, T. Miyamura, H. Utsumi, K. Inoue, M. Umeda, Membrane structure of the hepatitis B virus surface antigen particle. *J. Biochem.* 127 (2000) 543–550.
- [13] R.J. Gilbert, L. Beales, D. Blond, M.N. Simon, B.Y. Lin, F.V. Chisari, D.I. Stuart, D. J. Rowlands, Hepatitis B small surface antigen particles are octahedral. *Proc Natl Acad Sci U S A* 102 (2005) 14783–14788.
- [14] D. Diminsky, R. Schirmbeck, J. Reimann, Y. Barenholz, Comparison between hepatitis B surface antigen (HBsAg) particles derived from mammalian cells (CHO) and yeast cells (*Hansenula polymorpha*): composition, structure and immunogenicity. *Vaccine* 15 (1997) 637–647.
- [15] V. Bruss, D. Ganem, The role of envelope proteins in hepatitis B virus assembly. *Proc Natl Acad Sci U S A* 88 (1991) 1059–1063.
- [16] F. Gaviñanes, J.M. Gonzalez-Ros, D.L. Peterson, Structure of hepatitis B surface antigen. Characterization of the lipid components and their association with the viral proteins. *J. Biol. Chem.* 257 (1982) 7770–7777.
- [17] E. Guerrero, F. Gaviñanes, D.L. Peterson, Model for the protein arrangement in HBsAg particles based on physical and chemical studies. in: A. J. Zuckerman (Ed.), *Viral Hepatitis and Liver Disease*. Alan R Liss Inc., New York, 1988, pp. 606–613.
- [18] B.R. Lentz, Membrane 'fluidity' as detected by diphenylhexatriene probes. *Chem. Phys. Lipids* 50 (1989) 171.
- [19] B.R. Lentz, Use of fluorescent probes to monitor molecular order and motions within liposome bilayers. *Chem. Phys. Lipids* 64 (1993) 99–116.
- [20] G. Diaz, M. Melis, B. Batetta, F. Angius, A.M. Falchi, Hydrophobic characterization of intracellular lipids in situ by Nile Red red/yellow emission ratio. *Micron* 39 (2008) 819–824.
- [21] P. Greenspan, E.P. Mayer, S.D. Fowler, Nile Red: a selective fluorescent stain for intracellular lipid droplets. *J. Cell. Biol.* 100 (1985) 965–973.
- [22] A.P. Demchenko, Y. Mely, G. Duportail, A.S. Klymchenko, Monitoring biophysical properties of lipid membranes by environment-sensitive fluorescent probes. *Biophys. J.* 96 (2009) 3461–3470.
- [23] A.S. Klymchenko, Y. Mely, A.P. Demchenko, G. Duportail, Simultaneous probing of hydration and polarity of lipid bilayers with 3-hydroxyflavone fluorescent dyes. *Biochim. Biophys. Acta* 1665 (2004) 6–19.
- [24] T. Ozturk, A.S. Klymchenko, A. Capan, S. Oncul, S. Cikrikci, S. Taskiran, B. Tasan, F.B. Kaynak, S. Ozbey, A.P. Demchenko, New 3-hydroxyflavone derivatives for probing hydrophobic sites in microheterogeneous systems. *Tetrahedron* 63 (2007) 10290–10299.
- [25] A.S. Klymchenko, G. Duportail, T. Ozturk, V.G. Pivovarenko, Y. Mely, A. P. Demchenko, Novel two-band ratiometric fluorescence probes with different location and orientation in phospholipid membranes. *Chem. Biol.* 9 (2002) 1199–1208.
- [26] V.V. Shynkar, A.S. Klymchenko, G. Duportail, A.P. Demchenko, Y. Mely, Two-color fluorescent probes for imaging the dipole potential of cell plasma membranes. *Biochim. Biophys. Acta* 1712 (2005) 128–136.
- [27] J. Eisinger, G. Navon, Fluorescence quenching and isotope effect of tryptophan. *J. Chem. Phys.* 50 (1969) 2069–2077.
- [28] E. Bombarda, A. Ababou, C. Vuilleumier, D. Gerard, B.P. Roques, E. Piemont, Y. Mely, Time-resolved fluorescence investigation of the human immunodeficiency virus type 1 nucleocapsid protein: influence of the binding of nucleic acids. *Biophys. J.* 76 (1999) 1561–1570.
- [29] Y. Mely, N. Jullian, N. Morellet, H. De Rocquigny, C.Z. Dong, E. Piemont, B. P. Roques, D. Gerard, Spatial proximity of the HIV-1 nucleocapsid protein zinc fingers investigated by time-resolved fluorescence and fluorescence resonance energy transfer. *Biochemistry* 33 (1994) 12085–12091.
- [30] A.K. Livesey, J.C. Brochon, Analyzing the distribution of decay constants in pulse-fluorimetry using the maximum entropy method. *Biophys. J.* 52 (1987) 693–706.
- [31] J. Azoulay, J.P. Clamme, J.L. Darlix, B.P. Roques, Y. Mely, Destabilization of the HIV-1 complementary sequence of TAR by the nucleocapsid protein through activation of conformational fluctuations. *J. Mol. Biol.* 326 (2003) 691–700.

- [32] J.P. Clamme, J. Azoulay, Y. Mely, Monitoring of the formation and dissociation of polyethylenimine/DNA complexes by two photon fluorescence correlation spectroscopy. *Biophys. J.* 84 (2003) 1960–1968.
- [33] N.L. Thompson, Fluorescence correlation spectroscopy. in: J.R. Lakowicz (Ed.), *Topics in Fluorescence Spectroscopy*, vol. 1. Plenum Press, New York, 1991.
- [34] P. Didier, J. Godet, Y. Mely, Two-photon two-focus fluorescence correlation spectroscopy with a tunable distance between the excitation volumes. *J. Fluoresc.* 19 (2009) 561–565.
- [35] Q. Zhao, Y. Wang, D. Freed, T.M. Fu, J.A. Gimenez, R.D. Sitrin, M. W. Washabaugh, Maturation of recombinant hepatitis B virus surface antigen particles. *Hum. Vaccin.* 2 (2006) 174–180.
- [36] J.R. Lakowicz, *Principles of Fluorescence Spectroscopy*, second ed. Academic/Plenum Publishers, New York, 1999.
- [37] M.R. Eftink, Fluorescence quenching: theory and applications. in: J.R. Lakowicz (Ed.), *Topics in Fluorescence Spectroscopy*, vol. 2. Plenum Press, New York, 1991.
- [38] S.S. Lehrer, Solute perturbation of protein fluorescence. The quenching of the tryptophyl fluorescence of model compounds and of lysozyme by iodide ion. *Biochemistry* 10 (1971) 3254–3263.
- [39] A. Siemiarczuk, B.D. Wagner, W.R. Ware, Comparison of the maximum entropy and exponential series methods for the recovery of distributions of lifetimes from fluorescence lifetime data. *J. Phys. Chem.* 94 (1990) 1661–1666.
- [40] W. Stryjewski, Z. Wasylewski, The resolution of heterogeneous fluorescence of multityryptophan-containing proteins studied by a fluorescence-quenching method. *Eur. J. Biochem.* 158 (1986) 547–553.
- [41] H.J. Stirk, J.M. Thornton, C.R. Howard, A topological model for hepatitis B surface antigen. *Intervirology* 33 (1992) 148–158.
- [42] M. Villain, P.L. Jackson, M.K. Manion, W.J. Dong, Z. Su, G. Fassina, T.M. Johnson, T.T. Sakai, N.R. Krishna, J.E. Blalock, De novo design of peptides targeted to the EF hands of calmodulin. *J. Biol. Chem.* 275 (2000) 2676–2685.
- [43] C. Raja, J. Ferner, U. Dietrich, S. Avilov, D. Ficheux, J.L. Darlix, H. de Rocquigny, H. Schwalbe, Y. Mely, A tryptophan-rich hexapeptide inhibits nucleic acid destabilization chaperoned by the HIV-1 nucleocapsid protein. *Biochemistry* 45 (2006) 9254–9265.
- [44] G.R. Dreesman, F.B. Hollinger, J.R. Suriano, R.S. Fujioka, J.P. Brunschwig, J. L. Melnick, Biophysical and biochemical heterogeneity of purified hepatitis B antigen. *J. Virol.* 10 (1972) 469–476.
- [45] C.M. Mangold, R.E. Streeck, Mutational analysis of the cysteine residues in the hepatitis B virus small envelope protein. *J. Virol.* 67 (1993) 4588–4597.
- [46] M.P. Cooreman, G. Leroux-Roels, W.P. Paulij, Vaccine- and hepatitis B immune globulin-induced escape mutations of hepatitis B virus surface antigen. *J. Biomed. Sci.* 8 (2001) 237–247.
- [47] P.D. Moens, M.K. Helms, D.M. Jameson, Detection of tryptophan to tryptophan energy transfer in proteins. *Protein J.* 23 (2004) 79–83.
- [48] N. Sonveaux, D. Thines, J.M. Ruyschaert, Characterization of the HBsAg particle lipid membrane. *Res. Virol.* 146 (1995) 43–51.
- [49] J.P. Segrest, M.K. Jones, H. De Loof, N. Dashti, Structure of apolipoprotein B-100 in low density lipoproteins. *J. Lipid. Res.* 42 (2001) 1346–1367.
- [50] R. Prassl, P. Laggner, Molecular structure of low density lipoprotein: current status and future challenges. *Eur. Biophys. J.* 38 (2009) 145–158.
- [51] N.P. Morales, C. Charlemchoung, R. Luechapudiporn, P. Yamanont, S. Fucharoen, U. Chantharaksi, Lipid fluidity at different regions in LDL and HDL of beta-thalassemia/Hb E patients. *Biochem. Res. Commun.* 350 (2006) 698–703.
- [52] T. Hevonoja, M.O. Pentikainen, M.T. Hyvonen, P.T. Kovanen, M. Ala-Korpela, Structure of low density lipoprotein (LDL) particles: basis for understanding molecular changes in modified LDL. *Biochim. Biophys. Acta* 1488 (2000) 189–210.
- [53] Q.T. Li, L. Tilley, W.H. Sawyer, F. Looney, C.C. Curtain, Structure and dynamics of microemulsions which mimic the lipid phase of low-density lipoproteins. *Biochim. Biophys. Acta* 1042 (1990) 42–50.
- [54] J.B. Massey, H.J. Pownall, Surface properties of native human plasma lipoproteins and lipoprotein models. *Biophys. J.* 74 (1998) 869–878.
- [55] P. Greenspan, P. Lou, Spectrofluorometric studies of Nile red treated native and oxidized low density lipoprotein. *Int. J. Biochem.* 25 (1993) 987–991.
- [56] J.D. Morrisett, R.L. Jackson, A.M. Gotto Jr., Lipoproteins: structure and function. *Annu. Rev. Biochem.* 44 (1975) 183–207.
- [57] A.S. Klymchenko, G. Duportail, A.P. Demchenko, Y. Mely, Bimodal distribution and fluorescence response of environment-sensitive probes in lipid bilayers. *Biophys. J.* 86 (2004) 2929–2941.
- [58] E.G. Finer, R. Henry, R.B. Leslie, R.N. Robertson, NMR studies of pig low- and high-density serum lipoproteins. Molecular motions and morphology. *Biochim. Biophys. Acta* 380 (1975) 320–327.
- [59] L.P. Aggerbeck, D.L. Peterson, Electron microscopic and solution X-ray scattering observations on the structure of hepatitis B surface antigen. *Virology* 141 (1985) 155–161.

B. Publication 3 : "The structure of HBsAg particles is not modified upon their adsorption on aluminium hydroxide gel"



## The structure of HBsAg particles is not modified upon their adsorption on aluminium hydroxide gel<sup>☆</sup>

Vanille J. Greiner<sup>a</sup>, Frédéric Ronzon<sup>b</sup>, Eric Larquet<sup>c</sup>, Bernard Desbat<sup>d</sup>, Catherine Estèves<sup>b</sup>, Julie Bonvin<sup>b</sup>, Frédéric Gréco<sup>b</sup>, Catherine Manin<sup>b</sup>, Andrey S. Klymchenko<sup>a</sup>, Yves Mély<sup>a,\*</sup>

<sup>a</sup> Laboratoire de Biophotonique et Pharmacologie, UMR 7213 CNRS, Université de Strasbourg, Faculté de pharmacie, 74 route du Rhin, 67401 Illkirch, France

<sup>b</sup> Sanofi Pasteur, 1541 av. Marcel Mérieux, 69280 Marcy l'étoile, France

<sup>c</sup> Laboratoire d'Enzymologie et Biochimie Structurales (LEBS) UPR 3082 du CNRS, Campus CNRS, Bâtiment 34, Avenue de la Terrasse, F-91198 Gif-Sur-Yvette Cedex, France

<sup>d</sup> CBMN UMR5248, Université de Bordeaux, 2 rue Robert Escarpit, 33607 Pessac, France

### ARTICLE INFO

#### Article history:

Received 2 April 2012

Received in revised form 25 May 2012

Accepted 30 May 2012

Available online 15 June 2012

#### Keywords:

Hepatitis B surface antigen

Aluminium hydroxide adjuvant

Fluorescence spectroscopy

### ABSTRACT

Current Hepatitis B vaccines are based on recombinant Hepatitis B surface antigen (HBsAg) virus-like particles adsorbed on aluminium (Al) gel. These particles exhibit a lipoprotein-like structure with about 70 protein S molecules in association with various types of lipids. To determine whether the adsorption on Al gel affects HBsAg structure, we investigated the effect of adsorption and mild desorption processes on the protein and lipid parts of the particles, using various techniques. Electron microscopy showed that the size and morphology of native and desorbed HBsAg particles were comparable. Moreover, infrared and Raman spectroscopy revealed that the secondary structure of the S proteins was not affected by the adsorption/desorption process. Affinity measurements with Surface Plasmon Resonance showed no difference between native and desorbed HBsAg for HBsAg-specific RF-1 monoclonal antibody. Steady-state and time-resolved fluorescence data of the intrinsic fluorescence of the S proteins further indicated that the adsorption/desorption of HBsAg particles on Al gel did not modify the environment of the most emitting Trp residues, confirming that the conformation of the S proteins remains intact. Moreover, using environment-sensitive 3-hydroxyflavone probes, no significant changes of the lipid core and lipid membrane surface of the HBsAg particles were observed during the adsorption/desorption process. Finally, the ratio between lipids and proteins in the particles was found to be similar before and after the adsorption/desorption process. Taken together, our data show that adsorption on Al gel does not affect the structure of the HBsAg particles.

© 2012 Elsevier Ltd. All rights reserved.

### 1. Introduction

Hepatitis B virus (HBV) infection is one of the most common human diseases. Each year, over one million people die from HBV-related chronic liver diseases [1,2], explaining the importance of Hepatitis B vaccination. Currently, Hepatitis B vaccines are based on recombinant HBsAg particles produced in yeast cells [3,4] or mammalian cells [5–7] and adsorbed on aluminium (Al) gel. HBsAg particles are 22-nm spherical particles composed of host cell-derived lipids and up to three virus-encoded glycoproteins,

designated as large, medium and small (L, M, S). Compared to mammalian-derived HBsAg particles, yeast-derived particles, like the *Hansenula polymorpha*-derived HBsAg particles used in this study, contain only unglycosylated S proteins. This protein is highly hydrophobic [8] and in tight association with lipids in the HBsAg particles [6,7]. The lipid part of the HBsAg particles is mainly composed of phospholipids [4,9,10]. Lipids have been shown to be involved in the antigenic properties of HBsAg particles [11], stabilizing their structure and the protein conformation [4]. Several studies characterized the lipid/protein organization of the HBsAg particles [4,6,8,9,12–14] and we recently proposed a model for the HBsAg structure [15]. The particle is thought to exhibit a lipoprotein-like structure with an ordered and rather rigid lipid interface and a more hydrophobic and fluid inner core. About 70 proteins are included in each particle with a part protruding out [9,16] and another deeply inserted in the lipid core [15].

Al hydroxide or phosphate are the most common adjuvants for human and veterinary vaccines [17], used to potentiate immune responses, enhancing antigen uptake [18] and stimulating

**Abbreviations:** Al, aluminium; CD, circular dichroism; EM, electron microscopy; HBsAg, Hepatitis B surface antigen; HBV, Hepatitis B virus; IR, infrared; PLGA, poly (D,L)-lactide-co-glycolide acid; RAMFc, rabbit anti-mouse IgG Fcγ antibody; SPR, Surface Plasmon Resonance.

<sup>☆</sup> This work was supported by a grant from Sanofi Pasteur.

\* Corresponding author. Tel.: +33 3 90 24 42 63; fax: +33 3 90 24 43 13.

E-mail address: [yves.mely@unistra.fr](mailto:yves.mely@unistra.fr) (Y. Mély).



immune-competent cells [19,20]. Despite this widespread use, their properties and interactions with antigens have been poorly investigated. In vaccines against HBV, the adsorption of the HBsAg particles on the Al gel results from the binding of the phosphate groups of the phospholipids from HBsAg surface with the hydroxyl groups of the Al gel [21] through a ligand-exchange mechanism. When the vaccine is administered upon intramuscular or subcutaneous injection, a part of the antigens adsorbed to Al gel is expected to be eluted when it comes in contact with interstitial fluid [22,23].

The induction of a strong immune response is dependent on the integrity of the antigens. Since epitopes are conformational in nature, the structure of protein antigens on the surface of Al gel needs to be better characterized [24]. Recently, antigens after Al adsorption have been characterized by monitoring their conformation on Al hydroxide [25–27] or after desorption from the adjuvant surface [28]. Since antigens are often strongly bound to the adjuvant, drastic methods of elution using surfactants [29] or extreme pH changes [30] have been used but were found to lead to alteration of the desorbed antigens. Only recently, a soft method of desorption using competing phosphate anions has been developed [31].

For HBsAg particles, an alteration of their structure after adsorption on Al gel was described [27], but the adsorbed particles were only observed by immunoelectron microscopy after denaturing size exclusion chromatography and sodium dodecyl sulphate-polyacrylamide gel electrophoresis in reducing conditions.

In this study, we compared with various complementary techniques, the structure of native and adsorbed HBsAg particles with that of HBsAg particles desorbed with a soft method in non-denaturing conditions. Altogether, our results show that the structure of HBsAg particles is not modified by adsorption/desorption process.

## 2. Materials and methods

### 2.1. Materials

HBsAg virus-like particles were produced by Sanofi Pasteur in the recombinant yeast *Hansenula polymorpha* and obtained in a highly purified form after successive steps of fermentation, extraction and purification. Aluminium gel (Alhydrogel, Superfos) was purchased from Brenntag Biosector. RF-1 mAb was purchased from the NIBSC (National Institute for Biological Standards and Control). FN4 and PPZ8 dyes were synthesized as described elsewhere [32–34].

### 2.2. Adsorption/desorption process

The protocol of adsorption/desorption of HBsAg particles was derived from that of Egan et al. [31]. At room temperature, Al gel (6 mg Al/mL) was mixed with HBsAg particles (324 µg/mL) in phosphate buffer (30 mM) at pH 7.4. The sample was mixed gently by end-over-end rotation for 30 min to obtain complete adsorption. The sample was then centrifuged at 1000 rpm during 5 min to pellet the adsorbed material. In order to softly desorb the particles from the gel, 400 mM of phosphate buffer was added to the pellet. The sample was mixed during 1 h, and centrifuged at 3000 rpm during 5 min. HBsAg particles in the supernatant are referred to desorbed particles. The percent of HBsAg particles released after the desorption process was found to be about  $10 \pm 2\%$ , in excellent agreement with the original paper on this method [31].

### 2.3. Electron microscopy (EM)

For negative staining, HBsAg solutions (native and Al-desorbed) were diluted in Tris 10 mM, NaCl 150 mM buffer, pH 7.4 to a concentration of 20 µg/mL, deposited on a 400 mesh full carbon-coated

glow-discharged grid and stained with 2% uranyl acetate [16]. EM measurements were performed with a JEOL JEM 2100F transmission electron microscope, as previously described [16].

### 2.4. Infrared spectroscopy (IR)

HBsAg samples (native and Al-desorbed) were concentrated using capped centrifuge device (Amicon, Millipore) with a cut-off of 5 kDa up to 1.8 mg/mL in protein.

Fourier transform infrared (FTIR)-transmission spectra were recorded with a Vector 70 spectrophotometer (Bruker, Germany) equipped with a highly sensitive photovoltaic MCT detector cooled with liquid nitrogen. The spectrophotometer was continuously purged with dry N<sub>2</sub> to remove water vapour. In addition, an automatic water vapour correction was performed using the Opus software 5.5 version (Bruker, Germany). The Aqua Spec accessory (Bruker, Germany) consisting in a transmission cell with a path length of 6.8 µm between CaF<sub>2</sub> windows, has been used. Spectra were obtained from the collection of 120 scans per sample at a resolution of 4 cm<sup>-1</sup> and were corrected from buffer. Protein secondary structures were calculated using the Confochek system (Brucker). Second-derivative spectra were obtained using the Opus software 2.2 version with a nine-point Savitsky-Golay derivative function.

### 2.5. Circular Dichroism (CD)

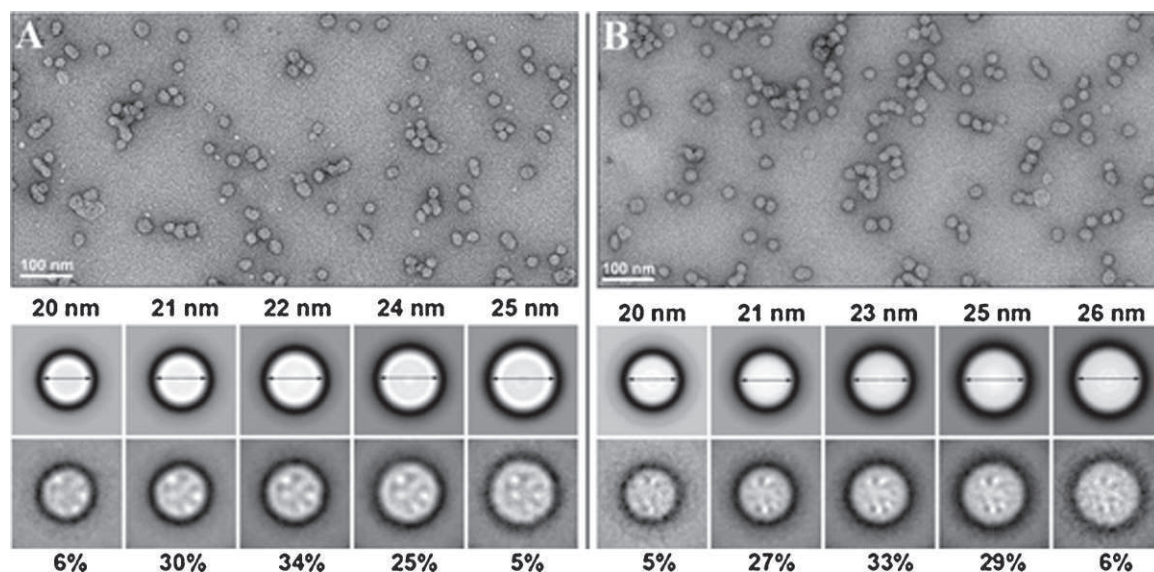
Far-UV CD spectra of native, adsorbed and desorbed HBsAg particles were recorded on a Jasco-810 spectropolarimeter in a 0.1 cm path-length circular cell (Hellma). To avoid sedimentation, the 0.2 mg/mL protein-containing samples were rotated at 52 rpm. Spectra were recorded at room temperature by averaging 3 scans from 280 to 180 nm (1 nm bandwidth) in 0.5 nm steps at a rate of 50 nm/min, and 1 s response. Blank spectra of either aqueous solutions or Al gel were used to correct the observed spectra. Data were analysed and smoothed by the means-movement method using the Jasco Spectra Analysis software. Spectra from 197 to 260 nm were subjected to secondary structure analysis using the Jasco CD Multivariate SSE software.

### 2.6. Surface Plasmon Resonance (SPR)

Antigen-antibody interactions were investigated with a temperature-controlled BIAcore 3000 instrument (Uppsala, Sweden). The running buffer contained 10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, and 0.005% (w/v) polysorbate 20. The flow rate was maintained at 70 µL/min, except during acid regeneration where the flow rate was 30 µL/min. Affinity-purified polyclonal rabbit anti-mouse IgG Fcγ antibody (RAMFc) was chemically immobilized on the CM5 sensor chip surface using an amine coupling kit (GE). The HBsAg-specific mAb RF-1 [35] was captured by RAMFc (80RU), and the affinity of HBsAg particles towards RF-1 mAb was determined by injecting different concentrations of HBsAg particles onto the sensor chip. The kinetic parameters and the equilibrium dissociation constant were determined using the BiaEval software.

### 2.7. Fluorescence spectroscopy

All experiments were performed in 8 mM phosphate buffer saline (PBS), NaCl 150 mM, pH 7.4, at 20 °C. Concentration of proteins for HBsAg particles was 0.4 µM, except for quantum yield measurements, where it was 2 µM. As the gel induces strong light scattering, time-resolved fluorescence decay, quantum yield, KI quenching and steady-state anisotropy could not be measured for the adsorbed particles. Characterization of the lipid and S proteins organization was performed, as previously



**Fig. 1.** Electron microscopy of negatively stained HBsAg particles. Native (panel A) and desorbed (panel B) HBsAg particles were negatively stained and imaged using EM. The lower panels correspond to the segmentation of a set of 1077 native particles (panel A) and 964 desorbed particles (panel B) determined using eigenvector in multivariate statistical analysis (MSA). The HBsAg particles appear as roughly spherical particles of about 23 nm diameter.

described [15]. Absorption spectra were recorded on a Cary 400 spectrophotometer. Both light scattering and gel absorbance were subtracted from the spectra of adsorbed particles. Fluorescence emission spectra and anisotropy were recorded at 20 °C on a FluoroMax spectrofluorometer (Jobin-Yvon) and a T-format SLM 8000 spectrofluorometer, respectively. Time-resolved fluorescence measurements were performed with the time-correlated, single-photon counting technique, as previously described [36,37].

### 3. Results

#### 3.1. Electron microscopy

The size of the native and desorbed HBsAg particles was determined by EM using negative staining methods. Selected fields were exposed to the electron beam under low dose conditions, using defocus values ranging from  $-1000$  nm to  $-1200$  nm. A set of 1077 native HBsAg particles and 964 desorbed HBsAg particles were extracted from ten numeric images. The size of native and desorbed particles was similar, ranging from 20 to 26 nm with a mean value of  $23 \pm 1$  nm (Fig. 1). HBsAg particles appeared as corrugated spherical particles but, probably due to the shadowing effect of the uranium salts and due to the heterogeneity of the surface of particles, no clear morphology details of the particles could be observed.

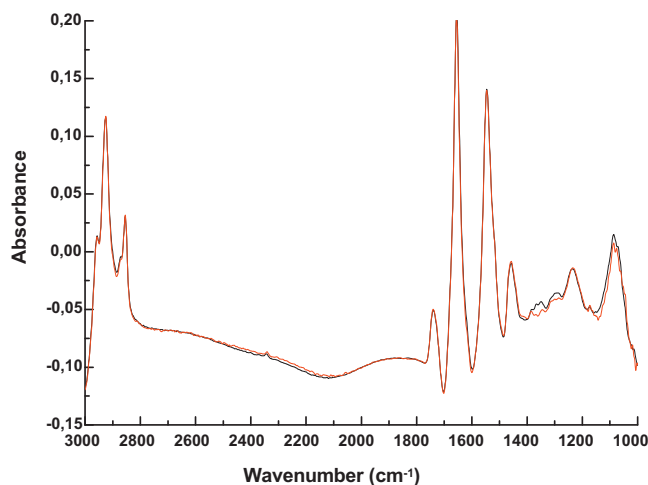
#### 3.2. Infrared, Raman and Circular Dichroism spectroscopy

The normalized IR spectra of native and desorbed HBsAg particles are totally superimposable (Fig. 2). In agreement with previous studies [8,38], the HBsAg particles show intense amide I and II bands centered at  $1655$   $\text{cm}^{-1}$  and  $1548$   $\text{cm}^{-1}$ , respectively, characteristic of proteins with a high content in  $\alpha$ -helix (Figs. 2 and S1). Curve fitting analysis of the IR spectra revealed that the secondary structure of S proteins in both native and Al-desorbed HBsAg particles contained about 58%  $\alpha$ -helix and negligible amount of  $\beta$ -sheet. Thus, no measurable alteration occurs in the secondary structure of the S proteins after desorption of HBsAg particles from the Al gel. Furthermore, the full overlap of the spectral regions ( $3000$ – $2800$   $\text{cm}^{-1}$ ) corresponding to lipids (Figs. 2 and S2) also indicates that the adsorption/desorption process does not modify

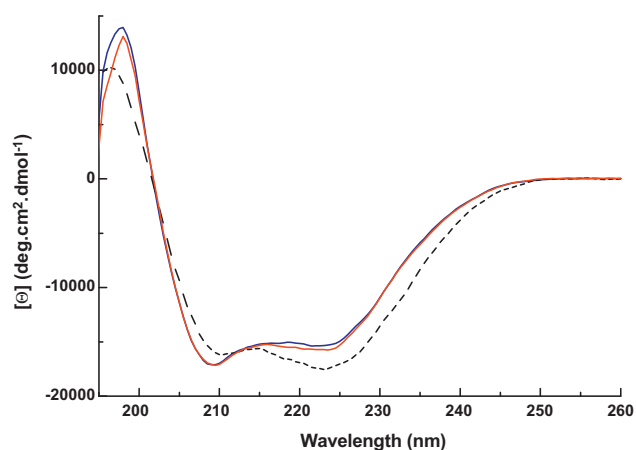
the lipid organization and the overall lipid:protein composition of HBsAg particles.

Far UV-CD spectra of native and desorbed HBsAg particles were also fully superimposable (Fig. 3), showing two strong negative bands at 208 and 222 nm, typical of protein  $\alpha$ -helical conformation [8,11,38]. In good agreement with Gavilanes and co-workers [11], the secondary structure contents estimated from these spectra were 47%  $\alpha$ -helix, 13%  $\beta$ -sheet, 10% turn and 30% random. Noticeably, the CD scan of Al-adsorbed HBsAg particles shows a slightly stronger signal at 222 nm and a slightly weaker signal at 196 nm than the spectra of native and Al-desorbed HBsAg particles, likely due to the corrections related to the strong light scattering of the Al gel. However, the percentages obtained for the secondary structure of S proteins are in the same range than for native particles, indicating that Al-adsorbed HBsAg particles retain their native structure.

Raman spectra (Fig. S3 and Table S1) confirmed that the helical and random coil structures are predominant in HBsAg particles and



**Fig. 2.** Normalized FTIR spectra of native (black line) and Al-desorbed (red line) HBsAg particles in the  $3000$ – $1000$   $\text{cm}^{-1}$  spectral region. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)



**Fig. 3.** Far UV-CD spectra of native (blue line), Al-adsorbed (dotted line) and Al-desorbed (red line) HBsAg particles. Spectra are blank subtracted. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

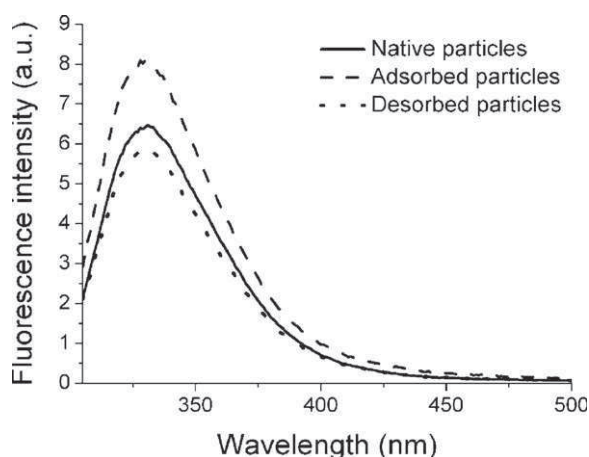
that the proportions of these structures do not change significantly after the adsorption/desorption process.

### 3.3. Affinity measurement of HBsAg towards RF-1 mAb by Surface Plasmon Resonance

Affinity measurements of HBsAg towards RF-1 mAb before and after desorption from Al gel was studied by SPR. RF-1 mAb is a neutralizing Ab that recognizes a conformational epitope of the S proteins [35]. Similar kinetic ( $k_{on}$ ,  $k_{off}$ ) and equilibrium ( $K_D$ ) binding constants (Table S2) were obtained for the interaction of RF-1 mAb with native and desorbed HBsAg particles. These results showed that no major modification occurs at the surface of HBsAg upon adsorption on Al gel.

### 3.4. Fluorescence spectroscopy

Fluorescence spectroscopy was used to assess changes in the tertiary structure of the S proteins in HBsAg particles. The maximum emission wavelength (331 nm) for the Al-adsorbed and Al-desorbed particles is comparable to that of the native particles (Fig. 4) and consistent with that reported previously [11,14,15], suggesting that the process of adsorption/desorption does not



**Fig. 4.** Comparison of the emission spectra of the native, Al-adsorbed and Al-desorbed HBsAg particles at pH 7.4. Excitation wavelength was 295 nm. The spectra have been corrected both for background fluorescence, Raman scattering and lamp fluctuations.

significantly modify the environment of the most emitting Trp residues in the S proteins. However, for the same concentration in particles, the maximum emission intensity for the Al-adsorbed particles is higher compared to that of the native particles, suggesting that the Al hydroxide gel could slightly enhance the fluorescence of the particles. The observed fluorescence enhancement by the Al hydroxide gel is fully consistent with the previously reported increase of the intrinsic fluorescence of several proteins bound to Al oxide nano-structured surfaces [39]. For native and Al-desorbed particles, a remarkable match of the quantum yields, time-resolved fluorescence parameters, bimolecular quenching constants  $k_q$  and steady-state anisotropy values was observed, indicating that the adsorption/desorption process has no influence on the structure and the conformation of the S-proteins (Table 1).

To further investigate the effect of adsorption/desorption process on the HBsAg particles, their lipid organization was compared using environment-sensitive 3-hydroxyflavone probes that bind to lipid-based structures. Due to excited state-intramolecular proton transfer reaction, these probes exhibit in the excited state, a normal ( $N^*$ ) and a tautomer ( $T^*$ ) form, which are differently sensitive to their environment, enabling to report on the local physicochemical properties of their binding site(s) [40,41]. While FN4 is a low polar probe that binds mainly to the lipid core of the HBsAg particles, the PPZ8 probe binds mainly to the lipid membrane surface [15]. Fluorescence spectra of both probes are comparable for native, Al-adsorbed and Al-desorbed particles (Fig. 5A and B), indicating that the structure of both the lipid core and surface of the HBsAg particles are not significantly altered by the adsorption/desorption process. Finally, from the fluorescence intensity ratio of the  $T^*$  band of the PPZ8 probe (proportional to the lipid quantity) relatively to the Trp emission (proportional to the protein quantity) (Fig. S4), no significant loss of lipids was found to accompany the adsorption/desorption process, in agreement with IR data.

## 4. Discussion

Due to the importance of antigen structure for vaccine immunogenicity, we characterized HBsAg particles by various techniques, before and during adsorption and after soft desorption on Al gel in order to determine whether the antigen structure is preserved.

The global structure of HBsAg particles was shown to remain intact after the soft desorption process since EM reveals similar morphology and size for desorbed and native HBsAg particles (Fig. 1). The measured size of HBsAg particles ( $\sim 23$  nm) is consistent with that of native HBsAg particles previously measured by fluorescence correlation spectroscopy (FCS) [15], atomic force microscopy (AFM) [14,16] and EM techniques [6,16].

IR and fluorescence spectroscopy techniques were previously shown to be good tools to investigate the secondary and tertiary protein structure of adsorbed antigens, and thus, detect conformational changes [25,26]. Application of these techniques, together with CD and Raman spectroscopy on HBsAg particles (Figs. 2–5, Table 1, Figs. S1–S3 and Table S1) demonstrates that the secondary structure and the conformation of the S proteins remain intact during and after adsorption on the Al gel. This was further confirmed by SPR data showing that the affinity of the surface S proteins for RF-1 mAb was not altered after the adsorption/desorption process (Table S2). A similar absence of changes in the structure of the S proteins was observed for HBsAg particles (before and after formulation) from vaccine formulation using poly (D,L)-lactide-co-glycolide acid (PLGA) microsphere as delivery system [38]. Since data from the literature showed that the protein structure can be either altered [25,26] or not [42] upon adsorption on the surface of Al-derived gels, the modification of the structure of adsorbed proteins appears to depend on the antigen type. In this respect, our



**Table 1**  
Steady-state and time-resolved fluorescence parameters of native and Al-desorbed HBsAg particles at pH 7.4.

Particles	$\tau_1^a$ (ns)	$\alpha_1$	$\tau_2$ (ns)	$\alpha_2$	$\tau_3$ (ns)	$\alpha_3$	$\tau_4$ (ns)	$\alpha_4$	$\langle\tau\rangle$ (ns)	$r^b$	$Q^c$	$k_q^d$ ( $M^{-1} s^{-1}$ )
Native particles	0.22	0.45	0.85	0.32	2.11	0.18	4.64	0.05	0.94	0.067	0.062	$2.0 \times 10^9$
Desorbed particles	0.20	0.44	0.85	0.33	2.21	0.18	5.01	0.04	0.99	0.065	0.061	$2.0 \times 10^9$
Native particles (data from Ref. [15])	0.31	0.31	0.85	0.44	2.20	0.20	4.40	0.05	1.13	0.067	0.058	$2.3 \times 10^9$

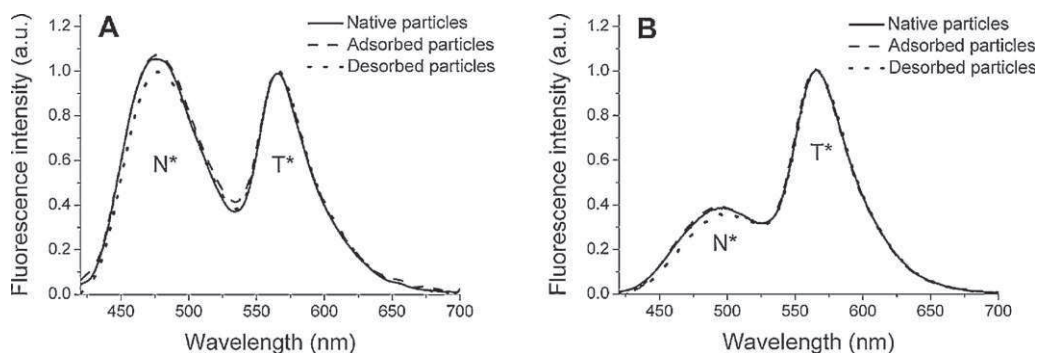
Excitation and emission wavelengths are 295 nm and 350 nm, respectively.

<sup>a</sup> The best fit to the time-resolved fluorescence decays was obtained using four lifetimes where  $\tau_1$  correspond to the fluorescence lifetimes,  $\alpha_1$  to the relative amplitudes, and  $\langle\tau\rangle$  to the mean lifetime.

<sup>b</sup>  $r$  correspond to the steady-state anisotropy value.

<sup>c</sup>  $Q$  correspond to the quantum yield.

<sup>d</sup>  $k_q$  correspond to the bimolecular quenching constant obtained with the external KI quencher that selectively quenches the solvent-exposed Trp residues.



**Fig. 5.** Fluorescence spectra of FN4 (A) and PPZ8 (B) probes in native, Al-adsorbed and Al-desorbed HBsAg particles at pH 7.4. Probe concentration was 0.6  $\mu$ M, while the lipid concentration for HBsAg particles was 6  $\mu$ M.

HBsAg particles seem to exhibit a more stable conformation than other types of antigen.

As HBsAg lipids are thought to preserve the overall structure of HBsAg particles and keep the protein S integrity, the effect of adsorption and desorption processes on the lipid part of the particles was also investigated. Adsorption through ligand-exchange of the phospholipids from the HBsAg particles with the hydroxyl groups from Al gel could damage the integrity of the external lipid layer of the particles and thus lead to a rearrangement of the S proteins. These modifications could be irreversible if part of the lipids stayed taped on the Al gel after desorption. Using IR (Figs. 2, S1 and S2) and fluorescence spectroscopy (Fig. S4), no change in the lipid:protein ratio was found to accompany the adsorption/desorption process, suggesting that the composition of the particles was not changed. Moreover, the PPZ8 and FN4 probes that bind to lipids show the same behaviour in native, adsorbed or desorbed HBsAg particles (Fig. 5). In addition, the peak positions of the lipid groups in the IR spectra were identical for native and Al-desorbed HBsAg particles. Thus, no changes seem to occur in the composition and organization of both the lipid core and surface of the particles. These data contrast with the IR spectra of HBsAg particles released from PLGA microspheres [38] that showed a partial delipidation.

Taken together, our data show that the structure of HBsAg particles remains intact during adsorption and desorption, in contrast with a previous report showing denaturation of the particles after desorption from the Al gel [27]. This discrepancy could be explained by the differences in the HBsAg particles and Al gel used, but is more probably related to the different protocols used to adsorb and desorb particles. In the present work, adsorption was performed in the presence of 30 mM phosphate that reduces the binding strength of the particles on the Al gel [21,31,43]. Due to the competing phosphate ions, the HBsAg particles can be desorbed in mild conditions that preserve their native structure. In contrast, in the previous report [27], harsher conditions, such as DTT and high temperature that could irreversibly modify the particles were used for desorption.

In conclusion, the combination of techniques used in this study revealed that adsorption of HBsAg particles on Al gel does not affect the lipid/protein organization, and thus, the antigen structure of the virus-like particles.

## Acknowledgements

The authors thank S. Lentsch-Graf and O. Adam for fruitful discussion, Diane Schneider and Audrey Rolland for technical assistance.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.vaccine.2012.05.082>.

## References

- [1] Ganem D, Varmus HE. The molecular biology of the hepatitis B viruses. *Annu Rev Biochem* 1987;56:651–93.
- [2] Robinson WS, Lutwick LI. The virus of hepatitis, type B (first of two parts). *N Engl J Med* 1976;295(November (21)):1168–75.
- [3] McAleer WJ, Buynak EB, Maigetter RZ, Wampler DE, Miller WJ, Hilleman MR. Human hepatitis B vaccine from recombinant yeast. *Nature* 1984;307(January (5947)):178–80.
- [4] Diminsky D, Schirmbeck R, Reimann J, Barenholz Y. Comparison between hepatitis B surface antigen (HBsAg) particles derived from mammalian cells (CHO) and yeast cells (*Hansenula polymorpha*): composition, structure and immunogenicity. *Vaccine* 1997;15(April–May (6–7)):637–47.
- [5] Diminsky D, Moav N, Gorecki M, Barenholz Y. Physical, chemical and immunological stability of CHO-derived hepatitis B surface antigen (HBsAg) particles. *Vaccine* 1999;18(August (1–2)):3–17.
- [6] Gilbert RJ, Beales L, Blond D, Simon MN, Lin BY, Chisari FV, et al. Hepatitis B small surface antigen particles are octahedral. *Proc Natl Acad Sci U S A* 2005;102(October (41)):14783–8.
- [7] Satoh O, Imai H, Yoneyama T, Miyamura T, Utsumi H, Inoue K, et al. Membrane structure of the hepatitis B virus surface antigen particle. *J Biochem* 2000;127(April (4)):543–50.
- [8] Guerrero E, Gavilanes F, Peterson DL. Model for the protein arrangement in HBsAg particles based on physical and chemical studies. In: Zuckerman AJ,

- editor. *Viral hepatitis and liver disease*. New York: Alan R Liss Inc.; 1988. p. 606–13.
- [9] Gavilanes F, Gonzalez-Ros JM, Peterson DL. Structure of hepatitis B surface antigen. Characterization of the lipid components and their association with the viral proteins. *J Biol Chem* 1982;257(July (13)):7770–7.
- [10] Satoh O, Umeda M, Imai H, Tunoo H, Inoue K. Lipid composition of hepatitis B virus surface antigen particles and the particle-producing human hepatoma cell lines. *J Lipid Res* 1990;31(July (7)):1293–300.
- [11] Gavilanes F, Gomez-Gutierrez J, Aracil M, Gonzalez-Ros JM, Ferragut JA, Guerrero E, et al. Hepatitis B surface antigen. Role of lipids in maintaining the structural and antigenic properties of protein components. *Biochem J* 1990;265(February (3)):857–64.
- [12] Aggerbeck LP, Peterson DL. Electron microscopic and solution X-ray scattering observations on the structure of hepatitis B surface antigen. *Virology* 1985;141(February (1)):155–61.
- [13] Stirk HJ, Thornton JM, Howard CR. A topological model for hepatitis B surface antigen. *Intervirology* 1992;33(3):148–58.
- [14] Zhao Q, Wang Y, Freed D, Fu TM, Gimenez JA, Sitrin RD, et al. Maturation of recombinant hepatitis B virus surface antigen particles. *Hum Vaccin* 2006;2(July–August (4)):174–80.
- [15] Greiner VJ, Egele C, Oncul S, Ronzon F, Manin C, Klymchenko A, et al. Characterization of the lipid and protein organization in HBsAg viral particles by steady-state and time-resolved fluorescence spectroscopy. *Biochimie* 2010;92(August (8)):994–1002.
- [16] Milhiet PE, Dossat P, Godefroy C, Le Grimellec C, Guigner JM, Larquet E, et al. Nanoscale topography of hepatitis B antigen particles by atomic force microscopy. *Biochimie* 2010;93(February (2)):254–9.
- [17] Gupta RK. Aluminum compounds as vaccine adjuvants. *Adv Drug Deliv Rev* 1998;32(July (3)):155–72.
- [18] Morefield GL, Sokolovska A, Jiang D, HogenEsch H, Robinson JP, Hem SL. Role of aluminum-containing adjuvants in antigen internalization by dendritic cells in vitro. *Vaccine* 2005;23(February (13)):1588–95.
- [19] Li H, Nookala S, Re F. Aluminum hydroxide adjuvants activate caspase-1 and induce IL-1beta and IL-18 release. *J Immunol* 2007;178(April (8)):5271–6.
- [20] Sokolovska A, Hem SL, HogenEsch H. Activation of dendritic cells and induction of CD4(+) T cell differentiation by aluminum-containing adjuvants. *Vaccine* 2007;25(June (23)):4575–85.
- [21] Iyer S, Robinett RS, HogenEsch H, Hem SL. Mechanism of adsorption of hepatitis B surface antigen by aluminum hydroxide adjuvant. *Vaccine* 2004;22(March (11–12)):1475–9.
- [22] Baier RE, Dutton RC. Initial events in interactions of blood with a foreign surface. *J Biomed Mater Res* 1969;3(March (1)):191–206.
- [23] Heimlich JM, Regnier FE, White JL, Hem SL. The in vitro displacement of adsorbed model antigens from aluminium-containing adjuvants by interstitial proteins. *Vaccine* 1999;17(July (22)):2873–81.
- [24] Hem SL, HogenEsch H, Middaugh CR, Volkin DB. Preformulation studies – the next advance in aluminum adjuvant-containing vaccines. *Vaccine* 2010;28(July (31)):4868–70.
- [25] Capelle MA, Brugger P, Arvinte T. Spectroscopic characterization of antibodies adsorbed to aluminium adjuvants: correlation with antibody vaccine immunogenicity. *Vaccine* 2005;23(February (14)):1686–94.
- [26] Jones LS, Peek LJ, Power J, Markham A, Yazzie B, Middaugh CR. Effects of adsorption to aluminum salt adjuvants on the structure and stability of model protein antigens. *J Biol Chem* 2005;280(April (14)):13406–14.
- [27] Tleugabulova D, Falcon V, Penton E. Evidence for the denaturation of recombinant hepatitis B surface antigen on aluminium hydroxide gel. *J Chromatogr B Biomed Sci Appl* 1998;720(December (1–2)):153–63.
- [28] Hutcheon CJ, Becker JO, Russell BA, Bariola PA, Peterson GJ, Stroop SD. Physicochemical and functional characterization of antigen proteins eluted from aluminum hydroxide adjuvant. *Vaccine* 2006;24(November (49–50)):7214–25.
- [29] Rinella JV, Workman RF, Hermodson MA, White JL, Hem SL. Elutability of proteins from aluminum-containing vaccine adjuvants by treatment with surfactants. *J Colloid Interface Sci* 1998;197(January (1)):48–56.
- [30] Rinella JV, White JL, Hem SL. Effect of pH on the elution of model antigens from aluminum-containing adjuvants. *J Colloid Interface Sci* 1998;205(September (1)):161–5.
- [31] Egan PM, Belfast MT, Gimenez JA, Sitrin RD, Mancinelli RJ. Relationship between tightness of binding and immunogenicity in an aluminum-containing adjuvant-adsorbed hepatitis B vaccine. *Vaccine* 2009;27(May (24)):3175–80.
- [32] Klymchenko AS, Duportail G, Ozturk T, Pivovarenko VG, Mely Y, Demchenko AP. Novel two-band ratiometric fluorescence probes with different location and orientation in phospholipid membranes. *Chem Biol* 2002;9(November (11)):1199–208.
- [33] Ozturk T, Klymchenko AS, Capan A, Oncul S, Cikrikci S, Taskiran S, et al. New 3-hydroxyflavone derivatives for probing hydrophobic sites in microheterogeneous systems. *Tetrahedron* 2007;63(41):10290–9.
- [34] Shynkar VV, Klymchenko AS, Duportail G, Demchenko AP, Mely Y. Two-color fluorescent probes for imaging the dipole potential of cell plasma membranes. *Biochim Biophys Acta* 2005;1712(June (2)):128–36.
- [35] Iwarson S, Tabor E, Thomas HC, Goodall A, Waters J, Snoy P, et al. Neutralization of hepatitis B virus infectivity by a murine monoclonal antibody: an experimental study in the chimpanzee. *J Med Virol* 1985;16(May (1)):89–96.
- [36] Bombarda E, Ababou A, Vuilleumier C, Gerard D, Roques BP, Piemont E, et al. Time-resolved fluorescence investigation of the human immunodeficiency virus type 1 nucleocapsid protein: influence of the binding of nucleic acids. *Biophys J* 1999;76(March (3)):1561–70.
- [37] Ramos P, Coste T, Piemont E, Lessinger JM, Bousquet JA, Chapus C, et al. Time-resolved fluorescence allows selective monitoring of Trp30 environmental changes in the seven-Trp-containing human pancreatic lipase. *Biochemistry* 2003;42(November (43)):12488–96.
- [38] Shi L, Caulfield MJ, Chern RT, Wilson RA, Sanyal G, Volkin DB. Pharmaceutical and immunological evaluation of a single-shot hepatitis B vaccine formulated with PLGA microspheres. *J Pharm Sci* 2002;91(April (4)):1019–35.
- [39] Ray K, Szmanski H, Lakowicz JR. Enhanced fluorescence of proteins and label-free bioassays using aluminum nanostructures. *Anal Chem* 2009;81(August (15)):6049–54.
- [40] Demchenko AP, Mely Y, Duportail G, Klymchenko AS. Monitoring biophysical properties of lipid membranes by environment-sensitive fluorescent probes. *Biophys J* 2009;96(May (9)):3461–70.
- [41] Klymchenko AS, Mely Y, Demchenko AP, Duportail G. Simultaneous probing of hydration and polarity of lipid bilayers with 3-hydroxyflavone fluorescent dyes. *Biochim Biophys Acta* 2004;1665(October (1–2)):6–19.
- [42] Dong A, Jones LS, Kerwin BA, Krishnan S, Carpenter JF. Secondary structures of proteins adsorbed onto aluminum hydroxide: infrared spectroscopic analysis of proteins from low solution concentrations. *Anal Biochem* 2006;351(April (2)):282–9.
- [43] Hansen B, Belfast M, Soung G, Song L, Egan PM, Capen R, et al. Effect of the strength of adsorption of hepatitis B surface antigen to aluminum hydroxide adjuvant on the immune response. *Vaccine* 2009;27(February (6)):888–92.

C. Publication 4 : "Characterization of the mechanisms of HIV-1 Vpr(52-96) internalization in cells"



## Research paper

Characterization of the mechanisms of HIV-1 Vpr(52–96) internalization in cells<sup>☆</sup>

Vanille J. Greiner<sup>a</sup>, Volodymyr Shvadchak<sup>a</sup>, Joëlle Fritz<sup>a</sup>, Youri Arntz<sup>a</sup>, Pascal Didier<sup>a</sup>, Benoît Frisch<sup>b</sup>, Christian Boudier<sup>a</sup>, Yves Mély<sup>a</sup>, Hugues de Rocquigny<sup>a,\*</sup>

<sup>a</sup>Laboratoire de Biophotonique et Pharmacologie, UMR 7213 CNRS, Université de Strasbourg, Faculté de Pharmacie, 74 Route du Rhin, 67401 Illkirch, France

<sup>b</sup>Laboratoire de Conception et Application de Molécules Bioactives, UMR 7199 CNRS, Université de Strasbourg, Faculté de Pharmacie, 74 route du Rhin, BP 60024, 67401 Illkirch, France

## ARTICLE INFO

## Article history:

Received 16 February 2011

Accepted 31 May 2011

Available online 21 June 2011

## Keywords:

Vpr

Confocal microscopy

Fluorescence

HIV

FLIM

## ABSTRACT

Addition of Vpr C-terminus to various cell types provokes cell apoptosis. This property was recently shown useful to develop inhibitors of cell proliferation. In that context, we investigated the cellular uptake of rhodamine- and fluorescein-labeled Vpr(52–96) peptides to understand the mechanism of Vpr C-terminus entry into cells. Dynamic light scattering data indicated that this peptide spontaneously formed polydispersed aggregates in cell culture medium. The fluorescently labeled Vpr(52–96) peptide was efficiently internalized, appearing either as large fluorescent patches in the cytoplasm or in a more diffuse form throughout the cell. Using isothermal titration calorimetry, we demonstrated that Vpr(52–96) can tightly associate with heparin, a glycosaminoglycan analog of heparan sulphate, suggesting a central role of the ubiquitous cell surface-associated heparan sulphate proteoglycans for the internalization of Vpr C-terminus. Fluorescently-labeled transferrin and methyl- $\beta$ -cyclodextrin showed that the Vpr C-terminus was mediated through clathrin- and caveolae/raft-dependent endocytosis. We found that Vpr C-terminus uptake was partly blocked at 4 °C suggesting the importance of membrane fluidity for Vpr C-terminus entry. In fact, atomic force microscopy and liposome leakage further indicated that the Vpr peptide can destabilize and disrupt model membrane bilayers, suggesting that this mechanism may contribute to the passive entry of the peptide. Finally, using fluorescence lifetime imaging, we found that the Vpr(52–96) peptide was stable in cells for at least 48 h, probably as a consequence of the poor accessibility of the peptide to proteolytic enzymes in aggregates.

© 2011 Elsevier Masson SAS. All rights reserved.

## 1. Introduction

The Vpr protein of HIV-1 plays a pivotal role in the virus pathophysiology by promoting the transactivation of the long terminal repeat (LTR) of HIV, and inducing both G2/M cell cycle arrest and apoptosis through interaction with numerous cellular proteins [1–3]. This regulatory protein is encapsidated in the nascent particle [4–6] and participates to the nuclear import of the HIV-1 pre-integration complex (PIC) in non-dividing cells [7–10].

NMR studies of Vpr and related peptides [11–15] established that Vpr contains three amphipathic  $\alpha$  helices spanning residues (17–33), (38–50) and (54–77) mutually oriented to generate a central hydrophobic plateau surrounded by two flexible N- and C-terminal domains [16]. This hydrophobic core is required for Vpr

oligomerization, which in turn promotes Vpr nuclear localization and Gag-Vpr recognition during the course of the viral assembly [17–20].

Some of Vpr functions such as NCp7 recognition, RNA binding and cell toxicity were shown to require the C-terminus of Vpr [21]. This C-terminus corresponds to a basic and soluble peptide able to enter cells causing apoptosis but the mechanism used by this peptide to enter in cells was poorly characterized [12,22–25]. Importantly, it was recently reported that addition of Vpr C-terminus to various tumor cell lines inhibits tumor cell proliferation [26,27]. In that context of new potent anticancer properties mediated by Vpr C-terminus we carefully re-investigated the uptake of this peptide in live cells.

Using confocal microscopy and two photon fluorescence lifetime imaging microscopy (FLIM), we found that both rhodamine- and fluorescein-labelled Vpr C-terminus peptides enter cells mostly through clathrin and caveolae/raft-mediated endocytosis. The peptide was shown to diffuse in the cytoplasm and to accumulate in vesicles remaining stable over 48 h. Moreover, we showed that Vpr(52–96) tightly binds to heparin and interacts with lipids,

<sup>☆</sup> This work was supported by ANRS (Association Nationale de la recherche sur le SIDA) and CNRS (Centre National de la Recherche Scientifique).

\* Corresponding author. Tel.: +33 3 90 24 41 03; fax: +33 3 90 24 43 12.

E-mail address: [hderocquigny@unistra.fr](mailto:hderocquigny@unistra.fr) (H. de Rocquigny).



causing membrane bilayer solubilization, a mechanism that could also explain Vpr C-terminus entry.

## 2. Experimental procedures

### 2.1. Materials

HMP resin *N,N*-diisopropylethylamine (DIEA) and *N*-methylpyrrolidone (NMP) were purchased from Appera and 9-fluorenylmethoxycarbonyl (Fmoc) amino acids from Neosystem (France). 5(6)-carboxyfluorescein (Fl) and 5(6)-carboxytetramethylrhodamine (Rh) were from Aldrich. Other reagents for peptide synthesis, including dimethylformamide (DMF), dichloromethane, piperidine and phospholipids (DOPC) were from SIGMA. Transferrin and Rhodamine-labelled DOPC were from Molecular Probes (Cergy Saint Christophe, France). The heparin used was a  $6.6 \pm 0.8$  kDa fragment purified using a published procedure [28] from a commercial preparation of porcine intestinal mucosa low molecular weight heparin (Calbiochem-VWR, Fontenay sous Bois, France).

### 2.2. Peptide synthesis and labeling

Peptides were synthesized using the Fmoc solid-phase method (0.1 mmole scale) with an automatic peptide synthesizer (ABI 433A, Appalar, France) [29] and ended with a caproic derivative. Then, 0.025 mole of Fmoc-deprotected peptidyl-resin was isolated and mixed with four equivalents of 5(6)-carboxyfluorescein (Fl) or 5(6)-carboxytetramethylrhodamine (Rh) combined with HBTU/HOBt and DIEA. This mixture was immediately added to 2 ml of NMP containing peptidyl-resin and stirred at 40 °C overnight. After filtration, the peptidyl-resin was cleaved for 2 h in a trifluoroacetic acid (TFA) solution containing the following scavenger: phenol (2%, w/v), thioanisole (5%, v/v) and ethanedithiol (2.5%, v/v). After precipitation, peptides were purified by reverse-phase high-performance liquid chromatography on a C8 column (uptionsphere 300 Å, 5 µm; 250 × 10, Interchim, France). Molecular masses obtained by ion spray mass spectrometry were: 5248.9, 5659 and 5605 for Vpr(52–96), Fl-Vpr(52–96) and Rh-Vpr(52–96) respectively, in agreement with the expected theoretical masses. Prior to use, peptides were dissolved in distilled water, aliquoted and stored at –20 °C. The sequence was <sup>52</sup>GDTWAGVEAIIIRILQQLFIHFRIGCRHSRIGVTQRRRANGASRS<sup>96</sup>.

### 2.3. Absorbance and fluorescence spectra

Absorbance and fluorescence spectra were recorded on a Cary 4000 UV–visible spectrophotometer (Varian) and FluoroMax3 spectrofluorimeter (Horiba, Jobin Yvon). Quantum yields were calculated using 5-carboxytetramethylrhodamine in EtOH (quantum yield,  $\phi = 0.95$ ) as a reference. Excitation wavelengths were 520 nm for Vpr-Rh and 470 nm for Vpr-Fl.

### 2.4. Cell culture and incubation with fluorescent peptides

HeLa cells ( $3 \times 10^5$ ) were cultured on 35 mm coverslips (µ-Dish IBIDI, Biovalley, France) in Dulbecco's modified Eagle medium supplemented with 10% fetal calf serum (Invitrogen Corporation, Cergy Pontoise, France) at 37 °C in a 5% CO<sub>2</sub> atmosphere. After 24 h, cells were incubated in Opti-MEM (Gibco), followed by addition of 1 µM for one hour of either Rh-Vpr(52–96) or Fl-Vpr(52–96). Then, cells were thoroughly washed with PBS (Phosphate Buffer Saline), incubated 30 s with 0.05 % trypsin, washed with PBS and further incubated with HBSS (Hanks' Balanced Salt Solution) for immediate

microscopy analysis (referred to as one hour post incubation) or incubated with D-MEM for an analysis 4 or 8 h post incubation.

To visualize the cellular uptake, cells were incubated either with 20 µg/ml of transferrin-fluorescein (60 min) or 5 mM of methyl-β-cyclodextrin (MβCD) (60 min) [30] or 0.25 mM *N,N*-dimethylammonium chloride (10 min), rinsed with Opti-MEM and then incubated with 1 µM Rh-Vpr(52–96). To inhibit the ATP-dependent endocytosis, HeLa cells were incubated for one hour at 4 °C with 1 µM Vpr(52–96). To quantify the location of Vpr in acidic compartments, 30 µM monensin (Sigma) was added three minutes prior to confocal microscopy imaging [31].

### 2.5. Confocal microscopy

Fluorescence confocal images of Fl- or Rh- tagged Vpr(52–96) peptides in living cells were taken using a confocal microscope (SPC UV1 AOBs, Leica) equipped with a HCX PL APO CS 63× oil immersion objective and an Ar/Kr laser. The fluorescein images were obtained by scanning the cells with a 488 nm laser line and filtering the emission with a 500–550 nm band-pass. For the rhodamine images, a 568 nm laser line was used in combination with a 580–700 nm band-pass filter.

### 2.6. Fluorescence lifetime imaging microscopy (FLIM)

FLIM measurements were performed using an in house constructed multi-photon laser-scanning microscope [19,32,33]. FLIM allows obtaining simultaneously images in intensity and images in which the contrast is given by the fluorescence lifetime. Two-photon excitation was obtained with a titanium/sapphire laser which delivers 100 fs pulses. For Rh, the excitation wavelength was 830 nm. Photons were collected using a two-photon short pass filter with a cut-off wavelength of 680 nm (F75-680, AHF, Germany), and a band-pass filter of  $520 \pm 17$  nm (F37-520, AHF, Germany). The fluorescence was directed to a fiber-coupled APD (SPCM-AQR-14-FC, Perkin Elmer), which was connected to a time-correlated single photon counting module (SPC830, Becker & Hickl, Germany), which operates in the reversed start-stop mode. Data were analyzed using a SPCImage V2.8 software (Becker & Hickl, Germany), which uses an iterative reconvolution method to recover the lifetimes from the fluorescence decays. For FLIM measurements, Rh-Vpr(52–96) was incubated at a 1 µM concentration with HeLa cells as described above. The resulting images were compared with those of cells incubated with 50 µM TAMRA.

### 2.7. Dynamic light scattering (DLS)

DLS experiments were performed with a Zetasizer Nano ZS equipped with a 4 mW He–Ne laser operating at 633 nm (Malvern Instruments). All measurements were performed at 25 °C at a measurement angle of 173°. Vpr(52–96) concentration was 1 µM. For each sample, three measurements were performed. Results were fitted with the “Dispersion Technology Software 5.03” (Malvern Instruments) to get the size distribution of the particles.

### 2.8. Isothermal titration calorimetry (ITC)

The heat effect for the interaction of heparin and Vpr(52–96) was analyzed at 20 °C in 50 mM Hepes, 0.1 M NaCl, pH 7.4 with a VP ITC microcalorimeter (Microcal, Northampton, MA, USA). Titration of heparin by Vpr(52–96) was performed by monitoring under constant stirring (310 rpm) the thermal power generated by repeated injections of 6 µl aliquots (in 12 s) of 129 µM Vpr(52–96) contained in the syringe into a 7.0 µM heparin solution contained in the 1.42 ml cell of the instrument. The total heat resulting from an

injection of titrant was calculated as the integral versus time of the experimental signal. A control experiment in which Vpr(52–96) was titrated into the buffer alone was done to determine the heat of dilution. Instrument control, data acquisition, and analysis were done with the VPViewer and Origin software provided by the manufacturer.

### 2.9. Large unilamellar vesicles of dioleoylphosphatidylcholine (LUV)

Large unilamellar vesicles of dioleoylphosphatidylcholine were obtained in 10 mM phosphate buffer, pH 7.4, by the classical extrusion method using an extruder (Lipex Biomembranes Inc) with polycarbonate filters of calibrated pores (Nucleopore) [34]. This generates monodisperse LUVs with a mean diameter of 0.11–0.12  $\mu\text{m}$  as measured with a Malvern Zetamaster 300 (Malvern, UK). For leakage assay, LUVs were prepared in the presence of carboxyfluorescein (CF) at a self-quenching concentration of 50 mM. Non-encapsulated CF was separated from the vesicle suspension through a Sephadex G-75 filtration column eluted with an iso-osmolar buffer containing 10 mM TRIS, 100 mM NaCl, at pH 7.4. For leakage kinetics, the lipids (20  $\mu\text{M}$ ) were treated with 0.5  $\mu\text{M}$  to 1  $\mu\text{M}$  of Vpr(52–96) in a quartz cuvette under constant stirring. The increase of fluorescence intensity was monitored during 20 min on the Fluoromax-3 spectrofluorimeter, with excitation at 480 nm (slit 1 nm bandwidth) and emission at 520 nm (slit 1 nm bandwidth).

### 2.10. Bilayer formation and atomic force microscopy (AFM)

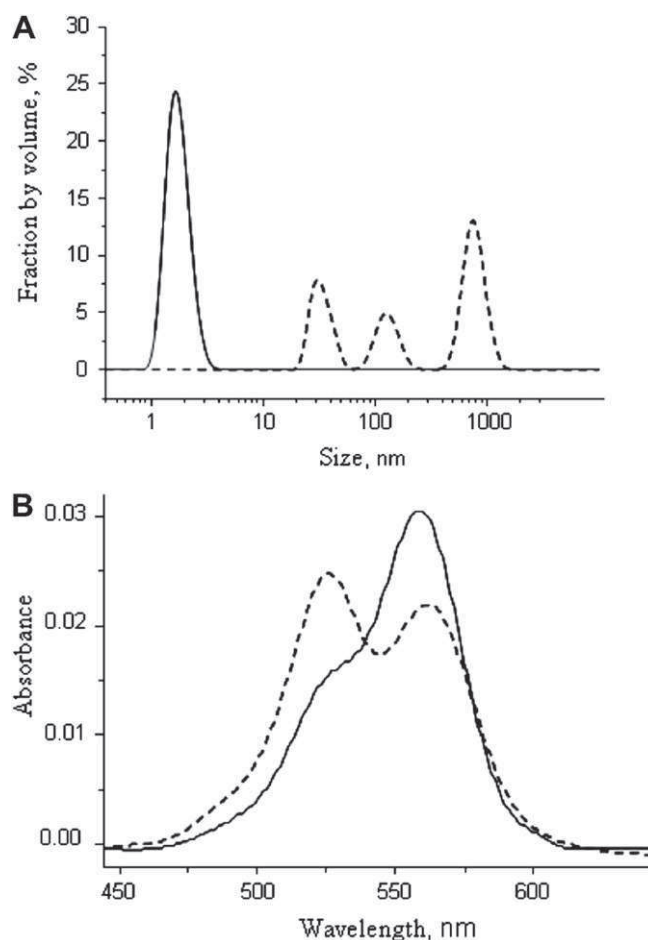
For AFM measurements, liposomes were prepared as described above and deposited on the mica surface. Both height and phase images were acquired in the tapping mode, using an AFM Solver Pro (Nt-MDT, Russia). Non-contact cantilevers NSG03 (NT-MDT) were used with a resonance frequency of 100 kHz in air and a spring constant of 1 N/m. The driving frequency of cantilevers was 28–32 kHz, the scan speed was typically 2–4 Hz, and the scan size was  $10 \times 10 \mu\text{m}$  with a resolution of  $512 \times 512$  points. After imaging the lipid bilayers, 100  $\mu\text{l}$  of the desired concentration of Vpr(52–96) was added to the bilayer and incubated at room temperature for 15 min.

## 3. Results

### 3.1. Vpr(52–96) forms oligomers in solution at physiological pH

It was previously shown that Vpr or Vpr C-terminus aggregate at pH 3 and that these aggregates are dissociated by TFE into small oligomers [12,13] or by acetonitrile into head-to-tail dimers [11]. To examine the possible oligomerisation of the Vpr(52–96) peptide in the conditions used for confocal microscopy (Opti-MEM, pH 7.4), this Vpr peptide was synthesized and its size distribution in the cell culture medium was compared by DLS to that with 30% TFE in water, pH 3. As shown in Fig. 1A, only a single population at 1.5–2 nm was observed in water/TFE consistent with Vpr(52–96) oligomers (solid line), while three populations of 30, 120 and 750 nm (dotted line) were evidenced in the cell culture medium. These three populations were also observed in PBS pH 7.4 (data not shown). This clearly indicated that the peptide forms heterogeneous aggregates in the cell culture medium.

In a next step, we examined this aggregation using the spectroscopic properties of rhodamine covalently bound to the N-terminus of the peptide (Rh-Vpr(52–96)). The absorption spectrum of 1  $\mu\text{M}$  Rh-Vpr(52–96) in water/TFE pH 3 provides a single absorption band with a maximum at 550 nm (Fig. 1B, solid line), similar to the spectrum of rhodamine in solution (data not shown). In sharp contrast,



**Fig. 1.** Size distribution of Vpr(52–96) and absorption properties of Rh-Vpr(52–96). A: DLS measurements were carried out on 1  $\mu\text{M}$  of Vpr(52–96) in Opti-MEM at pH 7.4 (dotted line) or in 30% TFE, pH 3 (solid line). All the measurements were recorded at 20  $^{\circ}\text{C}$ , fifteen minutes after dilution. B: UV spectra of 1  $\mu\text{M}$  Rh-Vpr(52–96) in Opti-MEM at pH 7.4 (dotted line) or in 30% TFE, pH 3 (solid line).

the absorption spectrum of the same concentration of Rh-Vpr(52–96) in Opti-MEM buffer shows a 5 nm red-shift of the Rh absorption maximum and a 20% decrease of its absorbance (Fig. 1B, dotted line), as well as the appearance of a blue-shifted peak at 520 nm. These dramatic changes in the absorption spectrum are typical of an excitonic interaction between the probes in their ground state and indicate that the fluorescent probes are a few angstroms apart in the Vpr(52–96) aggregates [35,36]. Excitonic interaction within the aggregates was confirmed by the dramatic drop in the fluorescence quantum yield of Rh-Vpr(52–96) from 0.45 in water/TFE to 0.005 in the Opti-MEM buffer [35,37]. To further understand the observed rhodamine quenching for Rh-Vpr(52–96), time-resolved fluorescence measurements were performed. In Opti-MEM buffer, Rh-Vpr(52–96) shows a bi-exponential decay, characterized by a very short ( $\tau_1 = 60$  ps) and a long ( $\tau_2 = 2.35$  ns) component (Table 1). The long component only marginally contributes to the decay (1%) and may correspond either to a small fraction of free labelled Vpr(52–96) peptide or to unquenched labelled peptides at the top of the aggregates. The major short component (99%), not observed for Rh in Opti-MEM buffer, likely originates from the emission of the quenched rhodamine residues involved in exciton interaction within the aggregates of Rh-Vpr(52–96).

The amplitude of the long decay component increases significantly ( $\alpha_2 = 17\%$ ) in water/TFE, in line with the presence of small

**Table 1**  
Time resolved fluorescence parameters of Rh-Vpr(52–96).

		$\tau_1$ (ns)	$\tau_2$ (ns)	$\alpha_1$	$\alpha_2$
Free Rh	Buffer, pH 7.4	1.59 ( $\pm 0.02$ )	/	1	/
	30% TFE, pH 3	2.67 ( $\pm 0.04$ )	/	1	/
Rh-Vpr(52–96)	Buffer, pH 7.4	0.06 ( $\pm 0.01$ )	2.35 ( $\pm 0.03$ )	0.99 ( $\pm 0.02$ )	0.01 ( $\pm 0.01$ )
	30% TFE, pH 3	0.54 ( $\pm 0.02$ )	2.86 ( $\pm 0.04$ )	0.83 ( $\pm 0.04$ )	0.17 ( $\pm 0.03$ )

The fluorescence lifetimes,  $\tau_i$  and the relative amplitudes,  $\alpha_i$ , are expressed as means  $\pm$  standard error of the mean for three experiments.

oligomers in these conditions [12,13]. Moreover, the value of the  $\tau_1$  component (0.54 ns) indicates that quenching of Rh fluorophores still occurs in the small oligomers of 1.5–2 nm evidenced by DLS (Fig. 1A, solid line).

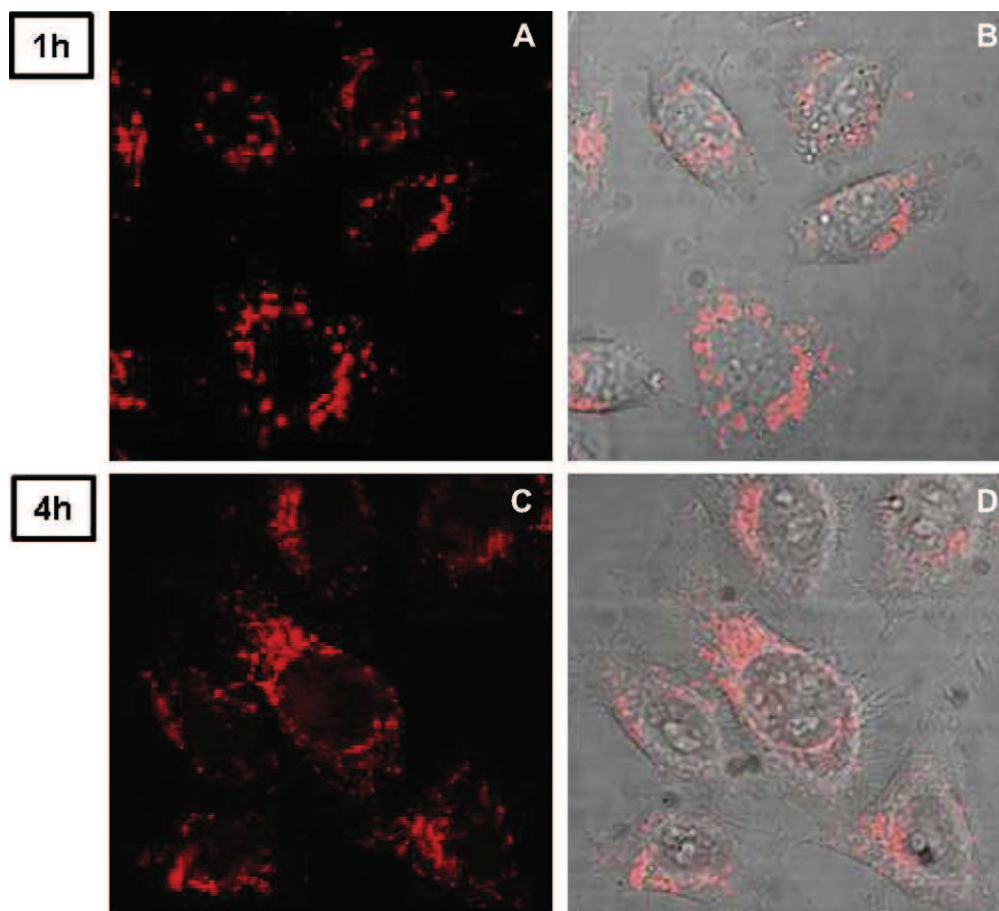
Thus, our data show that the Vpr(52–96) peptide forms large aggregates in the conditions used for confocal microscopy, which leads to dramatic changes in the spectroscopic properties of the fluorophore.

### 3.2. Vpr(52–96) localization in HeLa cells

To evaluate the mechanism of the cellular uptake of the Vpr(52–96) peptide, 1  $\mu$ M of the rhodamine labelled peptide was incubated for one hour with adherent HeLa cells. Then, cells were quickly washed with trypsin to remove the background [38] and left for one to four hours before image analysis. Propidium iodide exclusion revealed that this concentration of peptide does not induce cell toxicity while a clear cytotoxicity

was observed at 5  $\mu$ M, in line with the Vpr(52–96)-induced apoptosis effect (data not shown) [22,23]. All experiments were performed in plastic-bottom dishes allowing the live cells to be viewed by confocal microscopy in buffered solution, avoiding the possible artefacts in peptide distribution that could occur with fixed cells [38].

At one hour post incubation, large patches of fluorescence were observed in the cytosol (Fig. 2A and B). An increased fluorescence emission but no modification in the staining pattern was observed when the concentration of Rh-Vpr(52–96) was raised, in contrast to the concentration dependent uptake previously observed for the Ant peptide [39]. A detailed analysis of two-photon z-stack images revealed micrometric-sized patches at the surface of the cell but most of them were inside the cell. At four hours (Fig. 2C and D) or eight hours (data not shown) post-incubation, the fluorescent patches were more dispersed in the cytosol. Moreover, a diffuse fluorescence was also observed in the cytoplasm and to a lesser extent in the nucleus.



**Fig. 2.** Intracellular distribution of Rh-Vpr(52–96). HeLa cells were incubated with 1  $\mu$ M of Rh-Vpr(52–96) for one hour, washed and treated with trypsin. A and B: Cells were visualized by confocal microscopy one hour, after washing. The majority of the cells contains large aggregates of Vpr(52–96) dispersed in the cytoplasm. C and D: Cells were visualized four hours after the washing procedure. Note the punctuated distribution of the rhodamine labelled peptide and the faint diffusion of fluorescence in the nucleus.

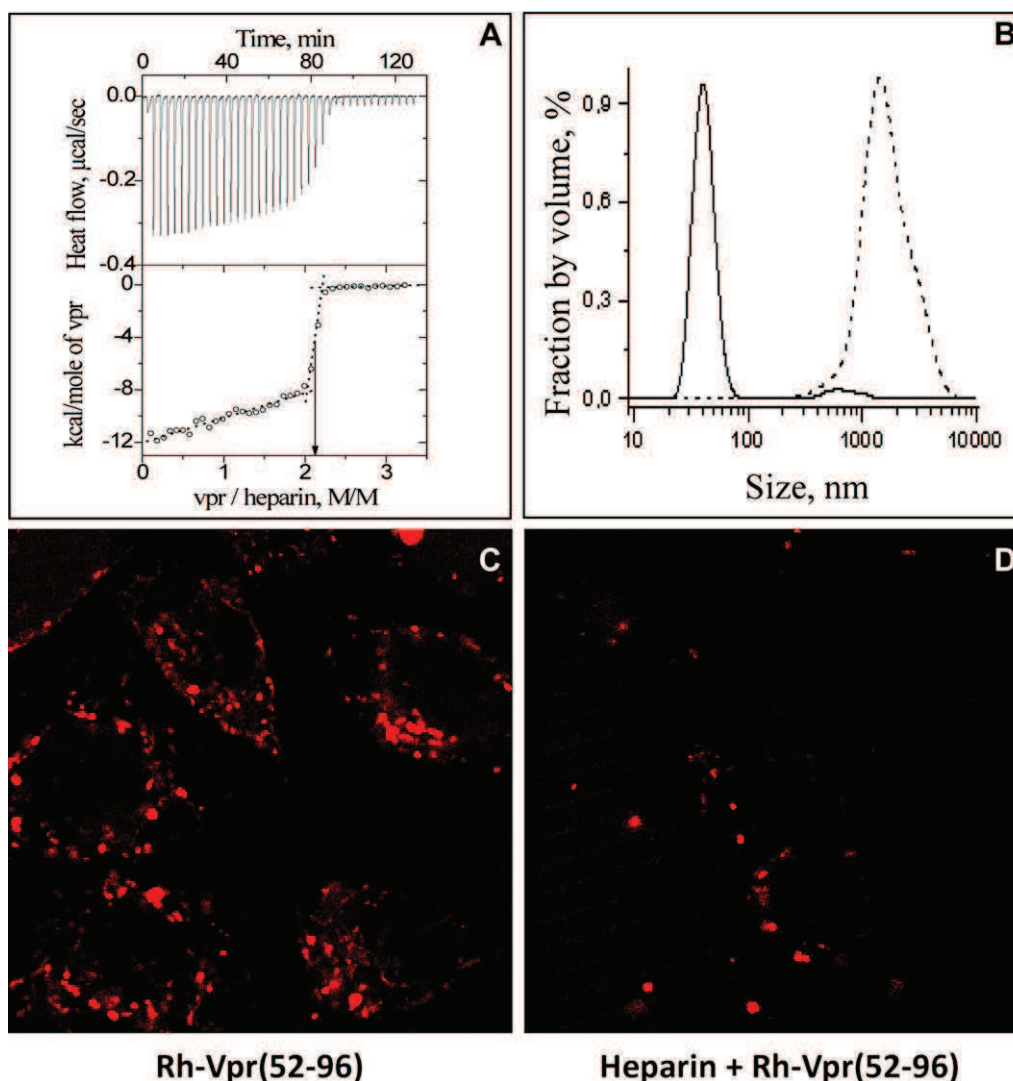


Thus, Vpr(52–96) penetrates HeLa cells, resulting in the presence of intracellular Vpr both in a diffuse form scattered all over the cell and in large cytoplasmic and perinuclear patches, consistent with Vpr(52–96) internalization in late endosomes and lysosomes.

### 3.3. Vpr(52–96) interacts strongly with heparin

Several studies demonstrated that cell surface heparan sulphate (HS), are directly involved in the CPP uptake (for review see [40]). To test whether these ubiquitously expressed HS proteoglycans could also be involved in Vpr C-terminus entry, we first investigated by ITC, the possibility for Vpr(52–96) to interact with heparin, a soluble glycosaminoglycan analog of HS [41]. Incremental addition of Vpr(52–96) to a fixed amount of heparin induced a substantial release of heat as illustrated in the upper panel of Fig. 3A by the negative peaks of experimental signal (the heat flow as a function of the time) characterizing an exothermic reaction. Integration of the peaks yielded the quantity of heat

released per injection of peptide solution into the reaction mixture. Fig. 3 (bottom panel) shows that the magnitude of the heat effect varied as heparin saturation progressed from an initial value of  $-11 \text{ kcal mol}^{-1}$  to about  $-6.7 \text{ kcal mol}^{-1}$  and then abruptly dropped to 0 in agreement with the formation of a tight complex between the peptide and the oligosaccharide. From the total heat released, we determined that the reaction enthalpy  $\Delta H_{\text{hep}}$  at  $20^\circ\text{C}$  was equal to  $-22 \pm 3 \text{ kcal mol}^{-1}$  per mole of complex formed. Moreover, a stoichiometry of about 2 Vpr(52–96) per mol of heparin fragment and a  $K_d \leq 10^{-8} \text{ M}$  were deduced from the binding curve as indicated by the arrow. A more precise determination of the binding parameters was not possible, due to the prone nature of Vpr(52–96) to aggregate [11–13] and to the high stability of its complex with heparin. Complementary information was obtained from DLS measurements showing that a mixture of  $1 \mu\text{M}$  Vpr(52–96) and  $10 \mu\text{M}$  heparin gave a single peak of micrometric-sized particles (Fig. 3B, dashed line), much larger than the about  $50 \text{ nm}$  size of free heparin (Fig. 3B, solid line). Compare to Fig. 1A,



**Fig. 3.** Interaction of Vpr(52–96) with heparin. A: Isothermal Titration Calorimetry of the interaction between heparin and Vpr(52–96) at  $20^\circ\text{C}$  in  $0.05 \text{ M}$  HEPES,  $0.1 \text{ M}$  NaCl pH 7.4. Panel 1: Thermal power generated upon incremental  $6 \mu\text{l}$  additions of a  $129 \mu\text{M}$  solution of peptide into the  $7.0 \mu\text{M}$  heparin solution in the reaction cell. Panel 2: Normalized quantities of heat calculated from the area of each peak as a function of the Vpr/heparin molar ratio. The data are corrected for the heat of dilution of the peptide solution. The arrow indicates the equivalence point. B: DLS measurements performed on  $10 \mu\text{M}$  of heparin diluted in DMEM buffer alone (solid line) or incubated 15 min with  $1 \mu\text{M}$  of Vpr(52–96) (dashed line). All the measurements were recorded at  $20^\circ\text{C}$ . C: control were  $1 \mu\text{M}$  of Rh-Vpr(52–96) was incubated with HeLa cells for one hour, treated and observed by confocal microscopy as described in Fig. 1. D:  $10 \mu\text{M}$  of heparin were added to HeLa cells in DMEM for 15 min prior to the addition of  $1 \mu\text{M}$  of Rh-Vpr(52–96). After one hour at  $37^\circ\text{C}$ , HeLa cells were washed and imaged.

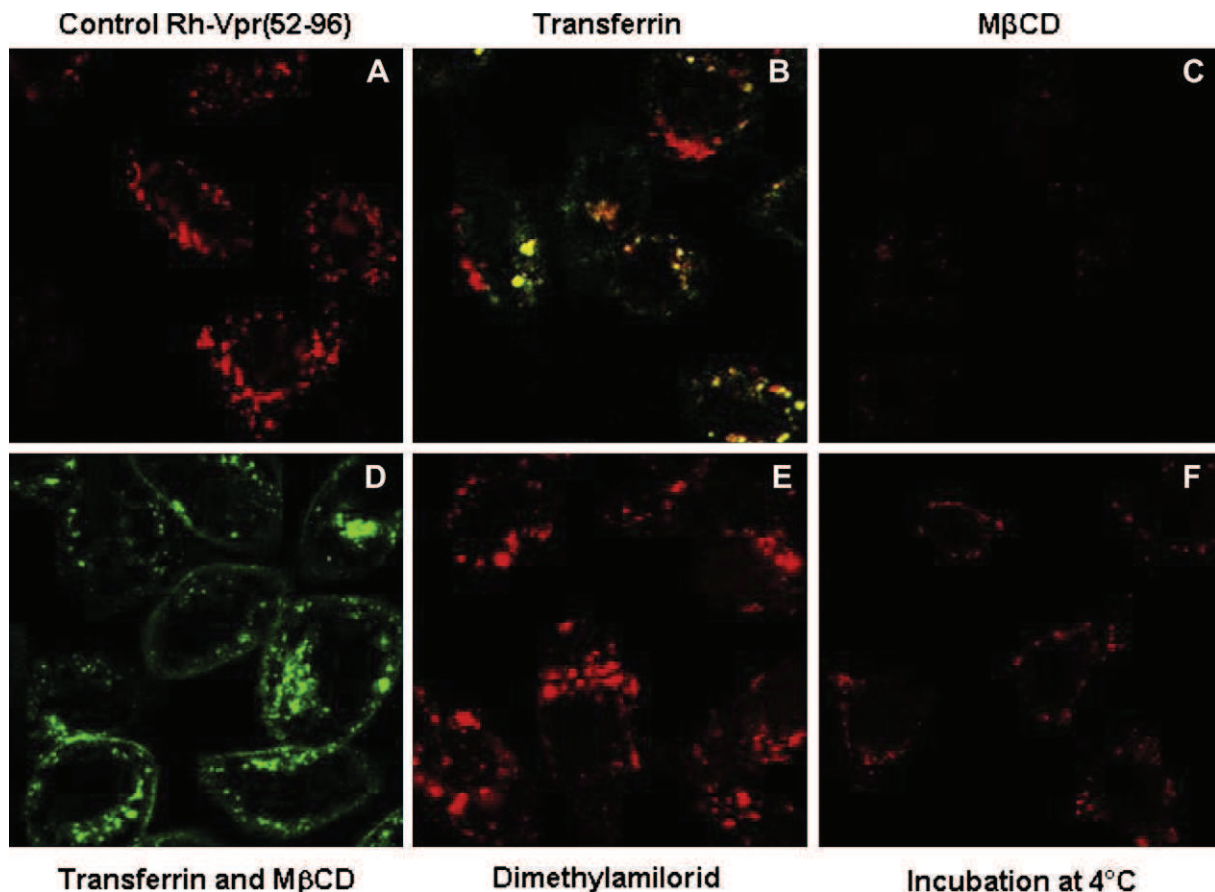
this result suggests that heparin mediates Vpr condensation. The high affinity of heparin for Vpr(52–96) strongly suggests that likely the membrane-attached HS also possess the capacity to associate with the peptide, a binding susceptible to trigger the internalization of this latter. To check this hypothesis, a competition experiment was carried out. Therefore, HeLa cells were treated 15 min with 10  $\mu$ M heparin followed by addition of 1  $\mu$ M Rh-Vpr(52–96). After the washing procedure, a drastic reduction of the fluorescent patches was observed within the cells (Fig. 3D) in comparison with non treated cells (Fig. 3C). This result indicates that upon association with extracellular heparin the peptide is mainly diverted from its binding sites on the cell membrane, suggesting that heparin likely efficiently competes with the membrane HS proteoglycans for the binding of Vpr(52–96). The interaction of this latter with HS appears thus as a prerequisite to its internalization.

#### 3.4. Different pathways for the cellular uptake of Vpr(52–96)

To characterize more in depth the mechanism of Vpr(52–96) cellular uptake, HeLa cells were incubated with Rh-Vpr(52–96) in different conditions. To test whether Vpr enters into cells by an energy-dependent endocytosis, we added both fluorescein labelled transferrin (Fl-Tr) and Rh-Vpr(52–96) to the cells. Transferrin is a protein that binds to its cell surface receptor and is internalized

through clathrin-associated endocytosis [42]. Superimposition of the green and red images (Fig. 4B) revealed the presence of Fl-transferrin-positive vesicles in green that do not contain Rh-Vpr(52–96), Rh-Vpr(52–96)-positive vesicles in red that do not contain Fl-transferrin and yellow coloured vesicles where the two proteins co-localize. The yellow vesicles suggest that clathrin-dependent endocytosis is partially responsible for the internalization of Vpr C-terminus.

To demonstrate that Vpr C-terminus can penetrate cell using clathrin-independent pathway, a cholesterol depletion assay was used [43]. Therefore, the fluorescent peptides were added to HeLa cells in the presence of methyl- $\beta$ -cyclodextrin (M $\beta$ CD). This compound depletes the cholesterol from the plasma membrane, disrupting the lipid rafts and caveolin-coated vesicles. With M $\beta$ CD, a large decrease of the intracellular fluorescence was observed (Fig. 4, compare A and C), indicating that internalization of the Vpr(52–96) peptide was drastically hampered. In contrast, M $\beta$ CD only moderately affected the internalization of transferrin (Fig. 4D), as previously reported [44]. Thus, cholesterol in the cytoplasmic membrane appeared essential for an efficient transduction of the Vpr(52–96) peptide. Interestingly, Coeytaux et al reported that cholesterol was only moderately involved in Vpr-mediated DNA transfection [45], suggesting different pathways for the free peptide and its complex with DNA, as reported for other CPPs [40].



**Fig. 4.** Cellular uptake of Rh-Vpr(52–96). A: HeLa cells were incubated with 1  $\mu$ M of Rh-Vpr(52–96) for one hour. Then cells were washed, treated with trypsin and visualized by confocal microscopy. B: HeLa cells were incubated with 1  $\mu$ M of Rh-Vpr(52–96) and 20  $\mu$ g/ml of Fl-transferrin for one hour, washed and treated with trypsin. This image corresponds to the merge of green and red fluorescent images scanned as described in Materials and Methods. Yellow dots show a co-localization of Rh-Vpr(52–96) and Fl-transferrin while red and green fluorescence indicate independent localization of the detected proteins. C: HeLa cells were incubated one hour with 5 mM methyl- $\beta$ -cyclodextrin (M $\beta$ CD) diluted in D-MEM, rinsed with Opti-MEM and then incubated with 1  $\mu$ M Rh-Vpr(52–96) for one hour. Then, cells were treated as described above. The weak intracellular staining suggests that M $\beta$ CD inhibits the Rh-Vpr(52–96) entry which is thus digested by trypsin. D: HeLa cells were treated by M $\beta$ CD as described above and incubated with 20  $\mu$ g/ml of Fl-transferrin. E: HeLa cells were incubated 10 min with 0.25 mM of dimethylamilorid diluted in D-MEM, rinsed with Opti-MEM and then incubated with 1  $\mu$ M Rh-Vpr(52–96) for one hour. The distribution pattern of Rh-Vpr(52–96) is close to that for the control. F: HeLa cells were incubated for one hour at 4  $^{\circ}$ C with 1  $\mu$ M of Rh-Vpr(52–96).

Since clathrin- and caveolin-coated vesicles are roughly 80–120 nm in diameter, micrometric-sized Vpr particles cannot be internalized through these vesicles. To test whether these large particles could enter through macropinocytosis, an actin-dependent endocytosis involved in the uptake of CPP-cargo molecules, CPP-modified liposomes [46] and HIV-1 virus [47], we incubated HeLa cells with the Rh-Vpr(52–96) peptide and dimethylamilorid, an inhibitor of macropinocytosis. Fig. 4E revealed a Vpr distribution pattern close to that in the absence of dimethylamilorid (Fig. 4A), suggesting that macropinocytosis plays only a marginal role in the internalization of the Vpr peptide.

Meanwhile, internalization through an energy-independent mechanism has also been reported for small peptides such as penetratin, (Arg)<sub>9</sub> and transportan which are efficiently internalized at both 37 °C and 4 °C (for review see [48]). To check whether this mechanism plays a role in the internalization of the Vpr(52–96) peptide, we incubated the rhodamine-labelled peptide with HeLa cells at 4 °C. As seen on Fig. 4F, cellular uptake of Vpr C-terminus was strongly decreased in this condition. However, intracellular fluorescent patches delineating the plasma membrane were observed suggesting that at least a part of Vpr(52–96) was internalized via an energy independent membrane crossing. Then the lack of spread of these patches in the cell could be explained by the temperature mediated effect on molecular diffusion and reduction of motor mediated transport along the cytoskeleton.

### 3.5. Vpr(52–96) binds to and destabilizes membrane bilayers

The effects of MβCD and of the temperature on Vpr C-terminus uptake are both linked to modifications of lipid membrane properties. To investigate the possible translocation of Vpr C-terminus into lipid bilayer, the interaction of Vpr(52–96) with supported DOPC bilayers was monitored by AFM, in 150 mM phosphate buffer, pH 7.4. In the absence of peptide, the DOPC bilayers appeared continuous and flat (Fig. 5A, image a), with a  $4.2 \pm 0.2$  nm thickness typical of a DOPC bilayer [49]. Bilayers were then incubated for

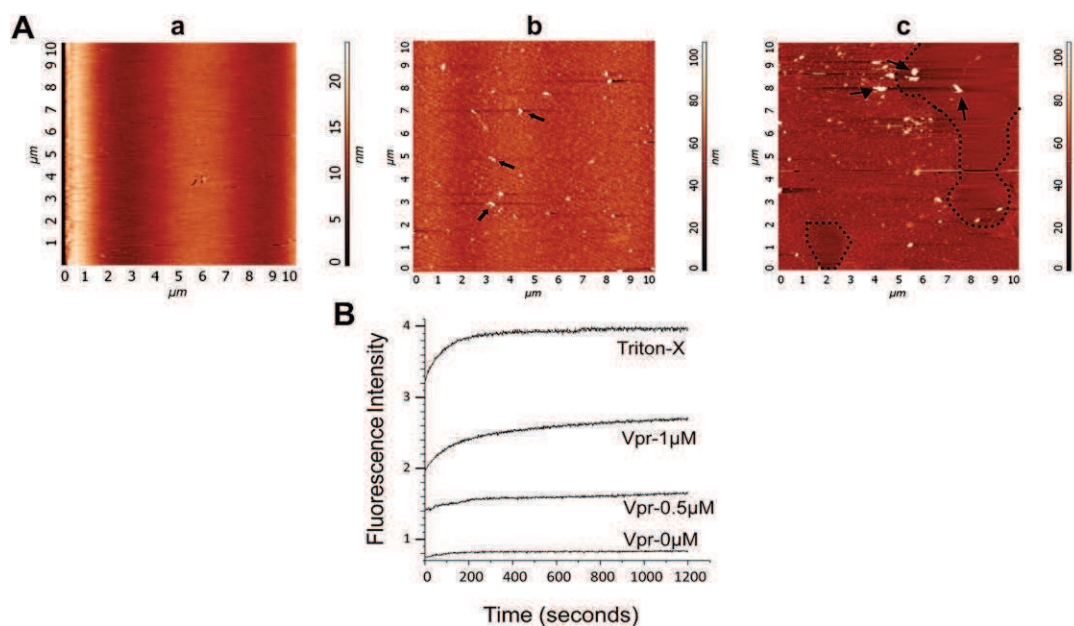
15 min with  $0.5 \mu\text{M}$  Vpr(52–96) at room temperature. AFM images revealed many small Vpr(52–96) particles at the bilayer surface with an average height and surface respectively of  $17.6 \pm 0.8$  nm and  $0.002 \mu\text{m}^2$  corresponding to an equivalent spherical diameter of  $47 \pm 4$  nm (Fig. 5A, image b). Moreover, in line with the polydispersity in the sizes of Vpr aggregates evidenced by DLS (Fig. 1), we also observed large particles with micrometer-scale diameter (arrows) bound to the bilayer. When the Vpr(52–96) concentration was increased to  $1 \mu\text{M}$ , similar particles were observed. In fact, the most striking and interesting difference with the lower concentration of Vpr is the reproducible appearance of large areas of uncoated mica surface (delineated by dotted line), suggesting that the Vpr(52–96) peptide can disrupt the surface bilayer, like a detergent (Fig. 5A, image c). In contrast, we found no evidence for fibril or channel formation with longer time of incubation or higher Vpr concentration.

To further characterize the ability of Vpr(52–96) to disrupt lipid bilayers, increasing concentrations of non-labeled peptide were added to fluorescein-containing DOPC liposomes. In the absence of peptide, the liposomes were weakly fluorescent due to the self-quenching effects of the high fluorescein concentration in the liposomes. The plot of fluorescence intensity versus time showed a 25% and 56% increase in the fluorescence intensity, respectively for  $0.5$  and  $1 \mu\text{M}$  of peptide (Fig. 5, B), indicating that the Vpr C-terminus promotes a concentration dependent leakage of DOPC vesicles. Similar data were also obtained with the negatively charged DOPS vesicles (data not shown).

Taken together, our data show that Vpr(52–96) can destabilize and disrupt lipid bilayers at micromolar concentration, a mechanism which may contribute to the internalization of this peptide into cells.

### 3.6. Vpr(52–96) accumulates in endosomes and lysosomes

Accumulation of Vpr(52–96) in punctuated vesicles likely corresponding to endosomal vesicles prompted us to further characterize



**Fig. 5.** Effect of Vpr(52–96) on membrane bilayers. A: AFM images of the interaction of DOPC bilayers with Vpr(52–96). Image a, control with DOPC bilayer on mica in phosphate buffer pH 7.4, NaCl 150 mM. Images b and c: interaction of  $0.5 \mu\text{M}$  and  $1 \mu\text{M}$  Vpr(52–96) with the DOPC bilayer. Vpr aggregates are indicated by black arrows. Note the appearance in image c of the mica surface resulting from Vpr-induced lipid desorption (dotted delineated surface). B: Kinetics of carboxy fluorescein leakage of DOPC vesicles ( $20 \mu\text{m}$ ) in TRIS buffer 10 mM, pH 7, 150 mM NaCl after addition of different concentrations of Vpr(52–96) and Triton-X100 (5% w/w). Excitation and emission wavelengths were 480 and 520 nm, respectively.



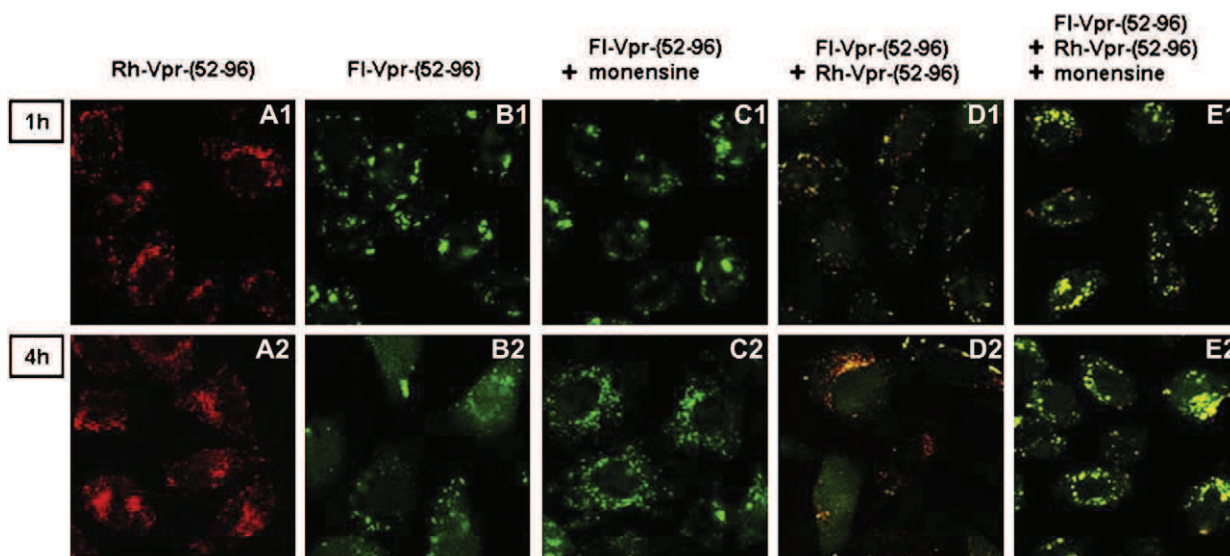
them. To this end, we compared the accumulation of Rh-labelled and FI-labelled Vpr(52–96), which exhibit respectively a pH-insensitive and pH-sensitive fluorescence [50]. Rh-Vpr(52–96) and FI-Vpr(52–96) were incubated at 1  $\mu$ M on HeLa cells and observed after one hour.

Both Rh-Vpr(52–96) (Fig. 6 A1) and FI-Vpr(52–96) (Fig. 6 B1) were observed to accumulate in large patches in the cytoplasm. The similar behavior of the two-labelled peptides suggests that the chromophore does not interfere with the intracellular localization of the Vpr(52–96) peptide. At 4 h post-incubation, a clear drop in the fluorescence of the FI-labelled vesicles (Fig. 6 B2) but not of the Rh-labelled (Fig. 6 A2) ones was observed. As a consequence of the decrease in the fluorescence of the FI-labelled vesicles, the low diffuse green fluorescent signal likely associated to cytoplasmic Vpr oligomers was much more visible (compare Fig. 6 B1 and B2). To check that this drop was a result of the quenching of FI by the acidic pH in vesicles, we recorded images after addition of 30  $\mu$ M monensin, a carboxylic ionophore that induces the neutralization of acidic intracellular compartments such as the lysosomes and acidic endosomes through  $H^+/Na^+$  exchange [51]. At 1 h post incubation, the localization pattern of FI-Vpr(52–96) in the presence of monensin was close to that in its absence (compare Fig. 6 B1 and C1). In sharp contrast, we observe large differences at 4 h post incubation (Fig. 6 B2 and C2). In the presence of monensin (Fig. 6 C2), a phenotype similar to that obtained with Rh-Vpr(52–96) was observed with a large number of fluorescent vesicles spread all over the cell excepted in the nucleus. To further demonstrate the localization of Vpr C-terminus in acidic compartment, equimolar concentration of Rh and FI peptides were pre-mixed and incubated on cells. Without monensin, superimposition of images recorded for the two channels gave rise to red coloured cells with some yellow dots suggesting that the fluorescein chromophore is quenched compared to the rhodamine (Fig. 6 D1 and D2). On the contrary, a yellow colour dominates on monensin treated cells in agreement with the dual emission of both fluorescein and rhodamine chromophores (Fig. 6 E1 and E2). Thus, our data strongly suggest that internalized Vpr peptides are routed to late endosomes and/or lysosomes.

### 3.7. Vpr(52–96) integrity within the cell

Due to its presence in both the cytoplasm and the endosomes, the Vpr(52–96) peptide could be the target for degradation both through ubiquitin-dependant/independent pathways in the cytoplasm and lysosomal digestion. This degradation likely explains the previously observed disappearance of DNA-Vpr complexes 24 h post transfection [45]. Many methods were proposed to evaluate the intracellular stability of CPPs [52]. Here, we took advantage of the fact that Vpr oligomerization clusters the rhodamine fluorophores, resulting in a decrease of their fluorescence lifetime (Table 1) due to self quenching and/or exciton coupling [53,54]. Accordingly, assuming that degradation of Vpr will affect its oligomerization and thus, induce an increase of the rhodamine lifetime, we imaged the cells incubated with Rh-Vpr(52–96) by FLIM as a function of time. Thus, changes of the fluorescence lifetime of Vpr aggregates will provide a direct evidence of the peptide stability within the cellular environment. These lifetimes were measured at each pixel or group of pixels and visualized by an arbitrary colour scale from blue to red.

HeLa cells were incubated with 1  $\mu$ M Rh-Vpr(52–96) and compared with cells incubated with 50  $\mu$ M rhodamine. This large concentration of free rhodamine was required as a consequence of its poor ability to enter into the cells. The free chromophore was found to be homogeneously distributed in the cells with an average fluorescence lifetime of 2.6–2.7 ns (Table 2), as evidenced by the yellow-red colour in Fig. 7A. When the chromophore is tethered to Vpr(52–96), a more punctuate staining is obtained at 4 h post incubation (Fig. 7B) in line with the confocal microscopy images (Fig. 2). Fig. 7B shows a heterogeneous colour distribution from green to blue corresponding to an average fluorescence lifetime value of 2.4 ns. A detailed analysis revealed that vesicles present a shorter lifetime around 2.1 ns. No significant modification of the fluorescence lifetimes was observed within the first 24–48 h (data not shown). In sharp contrast, at 72 h post incubation, two populations of vesicles were monitored with a fluorescent lifetime of 2.66 ns and 2.45 ns (Fig. 7D). The former value likely corresponds to free rhodamine and the latter to Rh-Vpr(52–96) containing



**Fig. 6.** Localization of Vpr(52–96) in late endosomes/lysosomes. Column A and B: HeLa cells were incubated respectively with 1  $\mu$ M Rh-Vpr(52–96) and 1  $\mu$ M of FI-Vpr(52–96) and observed at one and four hours post-incubation. Note the similarity of the distribution pattern suggesting that the chromophore does not influence the Vpr(52–96) behavior in cells. Column C: HeLa cells were incubated with 1  $\mu$ M FI-Vpr(52–96) as described for images B, and 30  $\mu$ M monensin were added just prior confocal imaging. Column D: HeLa cells were co-incubated with 0.5  $\mu$ M Rh-Vpr(52–96) and 0.5  $\mu$ M of FI-Vpr(52–96) and observed at one and four hours post-incubation. Column E: HeLa cells were co-incubated with 0.5  $\mu$ M Rh-Vpr(52–96) and 0.5  $\mu$ M of FI-Vpr(52–96) as described for images D, and 30  $\mu$ M monensin were added just prior confocal imaging.



**Table 2**

Lifetime values of free Rh and Rh-Vpr(52–96) in HeLa cells.

Compartment	Rhodamine (Rh)		Rh-Vpr(52–96)	
	Cytoplasm	Vesicles	Cytoplasm	Vesicles
4 h	2.63 ( $\pm 0.08$ )	2.67 ( $\pm 0.09$ )	2.4 ( $\pm 0.1$ )	2.1 ( $\pm 0.13$ )
72 h	2.7 ( $\pm 0.1$ )	2.65 ( $\pm 0.1$ )	2.66 ( $\pm 0.08$ )	2.45 ( $\pm 0.1$ )

The lifetime values were measured in the cytoplasm and intracytoplasmic vesicles. The values are average values resulting from measurements of 20 analyzed cells.

vesicles. Noticeably, in the control cells with free rhodamine, no change in the lifetime was observed after 72 h post incubation. These results suggest that Vpr is stable over 48 h and then is progressively degraded, notably in the cytoplasm.

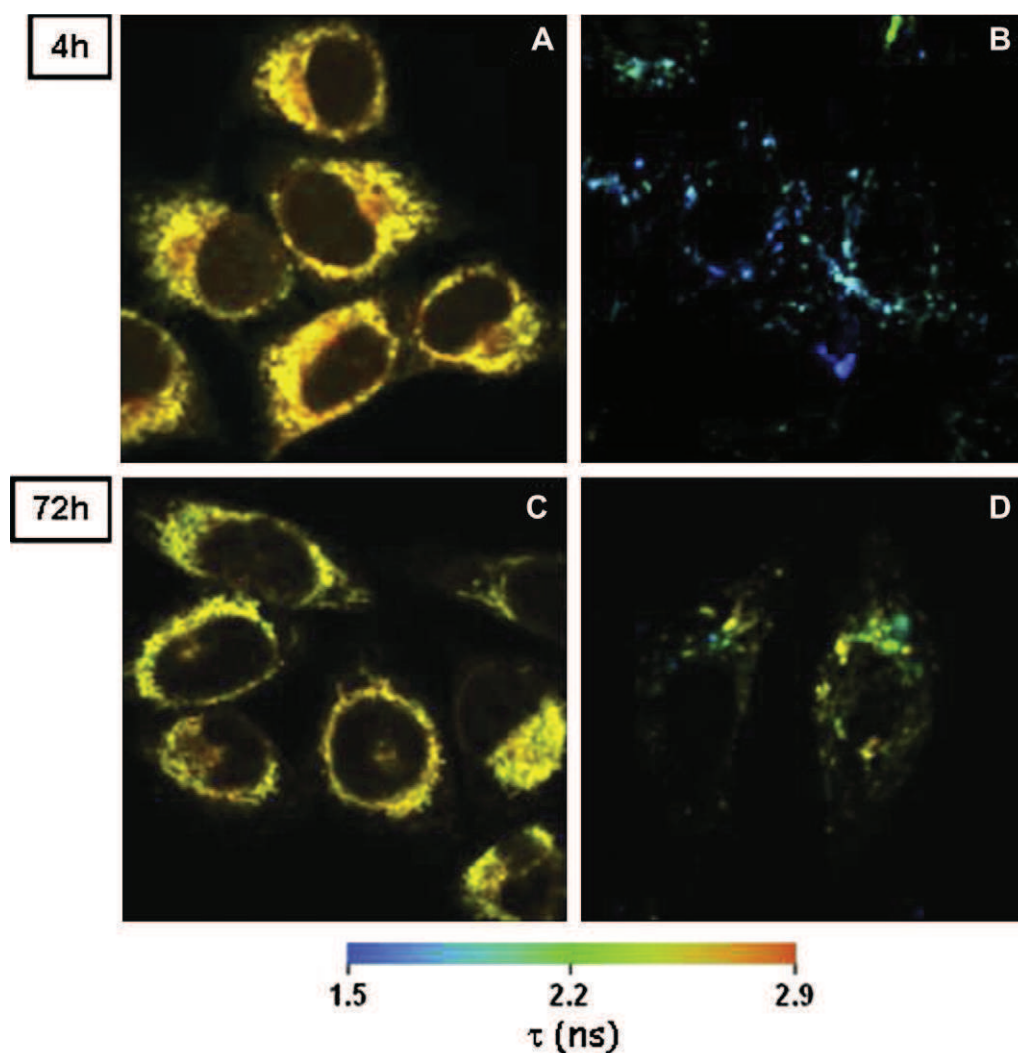
#### 4. Discussion

Recent results showed that the C terminal part of Vpr inhibit cell proliferation suggesting that this sequence could be used as potent

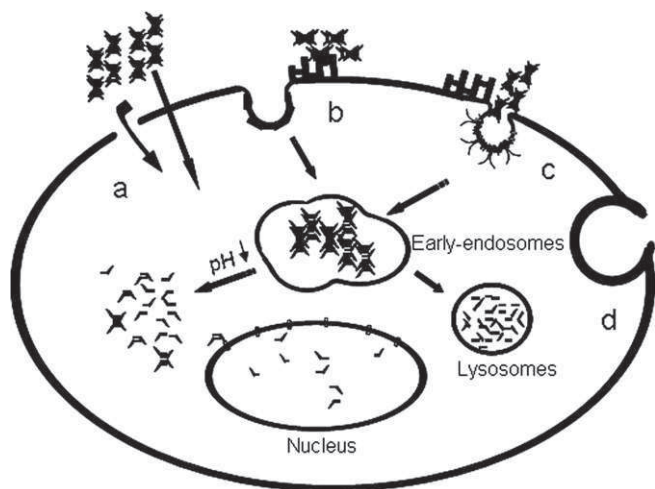
anti cancer agent [55]. To detail the mechanism of cell entry, we synthesized Vpr(52–96) peptides corresponding to the C-terminus of Vpr and labelled these peptides with either fluorescein or rhodamine probes at their N-terminus. This Vpr sequence was previously shown to efficiently enter into cells [7,22–24,56] even though contradictory results suggest the opposite [25]. This discrepancy could rely on the protocol used since in Sherman et al 2002, the Cy3 Vpr C-terminus entry was followed by flow cytometry which hardly distinguishes membrane-bound from internalized chromophore.

In this study, the labelled Vpr(52–96) peptides were found to be efficiently internalized in HeLa cells (Fig. 2) accumulating mainly in the cytoplasm both within heterogenous patches and in a more diffuse form. The patches probably result from accumulation of the aggregated peptides (Fig. 1) in vesicles while the diffuse staining could come from direct translocation of low molecular weight aggregates or from the exit of the peptide from acidic vesicles.

As depicted in Fig. 8, multiple uptake pathways are likely involved in the cellular entry of this peptide. The co-localization of the peptide with transferrin suggests that the clathrin mediated



**Fig. 7.** Intracellular stability of Rh-Vpr(52–96) aggregates. HeLa cells were incubated with 50  $\mu$ M rhodamine (image A) or 1  $\mu$ M Rh-Vpr(52–96) (image B). The fluorescence lifetimes were measured at 4 h post incubation and converted using a color scale ranging from blue (1.5 ns = short fluorescence lifetime) to red (3.0 ns = long fluorescence lifetime). A significant drop of the Rh-Vpr(52–96) fluorescence lifetime compared to free Rhodamine is evidenced by the blue colored patches. Images C and D correspond to lifetime images at 72 h of incubation. Note the decrease in the amount of short fluorescent lifetimes (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).



**Fig. 8.** Model depicting the Vpr(52–96) entry mechanism. a, b, c and d indicate respectively trans-membrane passage, clathrin-dependent endocytosis and caveolae-dependent endocytosis and macropinocytosis.

endocytotic pathway is involved in Vpr C-terminus entry (Fig. 8b). Nevertheless, this co-localization is partial (Fig. 4B) suggesting that alternative pathway could contribute to the Vpr C-terminus internalization. In fact, incubation of HeLa cell with methyl- $\beta$ -cyclodextrin, a drug that inhibits lipid raft-mediated endocytosis caused an almost complete inhibition of Vpr C-terminus uptake. In line with these results, a small part of Vpr C-terminus still enters the cell at 4 °C, suggesting that the loss of lipid membrane fluidity could interfere with Vpr C-terminus entry. Interestingly, this temperature effect appears more drastic for Tat transduction suggesting differences in the uptake pathways for these two peptides [48]. We also show that Vpr oligomers interact tightly with DOPC bilayers causing their disruption (Fig. 8a), depending on the peptide concentration (Fig. 5) while Tat interacts with liposome without lipid bilayer perturbation [57]. This Vpr-induced alteration of lipid bilayers is in line with the cell membrane permeabilization and herniation previously observed in the presence of Vpr [45,58,59]. In fact, the basic stretch of positively charged residues (77–96) of the Vpr peptide is probably capable to interact with the negatively charged phospholipids in the plasma membrane. This interaction likely promotes the insertion of the hydrophobic helix (amino acids 52–77), which in turn transiently destabilizes and/or disrupts the plasma membrane, enabling the diffusion of the aggregates into the cytoplasm [60]. Another model to explain Vpr entry would be the formation of pores, but these pores were never observed in cells [59,61].

We next evidenced that the Vpr C-terminus tightly associates with heparin, suggesting that the peptide is also capable to interact with the membrane-associated HS. The capacity of HS to bind numerous peptides such as the arginine rich region of Tat and to mediate or not their internalization is well documented [41,62,63]. As reported for Tat, this interaction induced the formation of aggregates on the cell surface and *in vitro* (Fig. 3) [64]. In this report we found that one molecule of the 6.6 kDa heparin fragment used can bind two molecules of peptide with an equilibrium dissociation constant  $K_d \leq 10^{-8}$  M. This value is at about one-two orders of magnitude smaller than those reported for the interaction between heparin and various CPPs [41,65]. This high affinity also explains that at a concentration as low as 100 nM of Vpr C-terminus efficiently penetrates the cell (data not shown). The binding of Tat to heparin/HS was reported to be mainly driven by electrostatic interactions [65]. Nevertheless, despite the fact that both Vpr(52–96) and Tat(47–57) possess 8 positive charges the affinity of the former

peptide for heparin is by far the highest, a property suggesting a contribution of non ionic interactions to complex formation. Interestingly, seven basic residues of Vpr(52–96) are located in the structurally flexible C-terminal region of the peptide comprising amino-acids 73–96 [11] which constitutes a cationic tail while residues 52–73 contains hydrophobic residues that adopt an  $\alpha$  helix conformation [16].

The diffuse fluorescence observed within the cells can be reasonably associated with Vpr oligomers. Taking into account that Vpr aggregates are dissociated into small oligomers in acidic conditions [11,12], this diffuse staining can result from the acidification of the endosomal vesicles, a process involved in the liberation of basic CPPs into the cytoplasm [66–68]. Alternatively, as mentioned above, Vpr oligomers may also cross directly the membrane bilayer, and diffuse freely into the cytoplasm.

The stability of CPPs with time in cells is poorly studied and probably dependent on the ubiquitin-proteasome pathway. The metabolic degradation of Tat (47–57), penetratin and hCT in cells was shown to occur with a half time value ranging from less than one hour to nine hours, depending on the peptide and cells conditions [69]. The Vpr(52–96) peptide was found to be stable for more than 48 h in cells, likely as a result of its poor accessibility to proteolytic enzymes in its oligomeric or aggregated form.

The mechanism of Vpr(52–96) peptide entry can be compared to that described for other CPPs. By analogy to Anp, Tat and oligoarginine peptides, the Vpr peptide appears to interact with proteoglycans, inducing its cellular uptake through either clathrin and caveolin/lipid raft mediated endocytosis. In contrast, Anp, Tat and oligoarginine peptides translocate cells with minimal perturbation of the membrane integrity [39,70–74] while Vpr(52–96) likely promotes membrane destabilization or disruption [58]. Another difference between Vpr(52–96) and other CPPs is the low contribution of macropinocytosis (Fig. 8d) in the intracellular entry of Vpr(52–96) in spite of the size of its aggregates [40]. At least, transduction capabilities of Vpr C-terminus could be compared with a 48 amino acid residues from the Epstein-Barr virus Zebra trans-activator [75]. This peptide share common feature with Vpr C-terminus such as a seven positive charged tail followed by a stretch of hydrophobic residues forming a leucine zipper domain. Nevertheless, this peptide was shown to enter cell mainly through a direct translocation pathway without significant energy dependent endocytosis.

In conclusion, this investigation of the cellular uptake of the Vpr C-terminus showed that both non-endocytotic and endocytotic uptake pathways are involved in its cellular internalization.

## Acknowledgments

Thanks to Andrey Klymchenko and to Sophie Martin for helpful discussions.

## List of abbreviations

Vpr	viral protein R
CPP	cell penetrating peptides
Tat	<i>trans</i> -acting activator of transcription
M $\beta$ CD	methyl- $\beta$ -cyclodextrin
TAMRA	carboxytetramethylrhodamine
Rh-Vpr(52–96)	N substitution of the Vpr C-terminus by carboxytetramethylrhodamine
Fl-Vpr(52–96)	N substitution of the Vpr C-terminus by fluorescein
FLIM	fluorescence lifetime imaging
HS	heparan sulfate
ITC	isothermal titration calorimetry
DLS	dynamic light scattering

## References

- [1] J.L. Andersen, V. Planelles, The role of Vpr in HIV-1 pathogenesis, *Curr. HIV Res.* 3 (2005) 43–51.
- [2] E. Le Rouzic, S. Benichou, The Vpr protein from HIV-1: distinct roles along the viral life cycle, *Retrovirology* 2 (2005) 11.
- [3] V. Planelles, S. Benichou, Vpr and its interactions with cellular proteins, *Curr. Top. Microbiol. Immunol.* 339 (2009) 177–200.
- [4] F. Bachand, X.J. Yao, M. Hrimech, N. Rougeau, E.A. Cohen, Incorporation of Vpr into human immunodeficiency virus type 1 requires a direct interaction with the p6 domain of the p55 gag precursor, *J. Biol. Chem.* 274 (1999) 9083–9091.
- [5] Y.L. Lu, R.P. Bennett, J.W. Wills, R. Gorelick, L. Ratner, A leucine triplet repeat sequence (LXX)4 in p6gag is important for Vpr incorporation into human immunodeficiency virus type 1 particles, *J. Virol.* 69 (1995) 6873–6879.
- [6] W. Paxton, R.I. Connor, N.R. Landau, Incorporation of Vpr into human immunodeficiency virus type 1 virions: requirement for the p6 region of gag and mutational analysis, *J. Virol.* 67 (1993) 7229–7237.
- [7] E. Le Rouzic, A. Mousnier, C. Rustum, F. Stutz, E. Hallberg, C. Dargemont, S. Benichou, Docking of HIV-1 Vpr to the nuclear envelope is mediated by the interaction with the nucleoporin hCG1, *J. Biol. Chem.* 277 (2002) 45091–45098.
- [8] R.A. Fouchier, B.E. Meyer, J.H. Simon, U. Fischer, A.V. Albright, F. Gonzalez-Scarano, M.H. Malim, Interaction of the human immunodeficiency virus type 1 Vpr protein with the nuclear pore complex, *J. Virol.* 72 (1998) 6004–6013.
- [9] M.G. Waldhuber, M. Bateson, J. Tan, A.L. Greenway, D.A. McPhee, Studies with GFP-Vpr fusion proteins: induction of apoptosis but ablation of cell-cycle arrest despite nuclear membrane or nuclear localization, *Virology* 313 (2003) 91–104.
- [10] T. Suzuki, N. Yamamoto, M. Nonaka, Y. Hashimoto, G. Matsuda, S.N. Takeshima, M. Matsuyama, T. Igarashi, T. Miura, R. Tanaka, S. Kato, Y. Aida, Inhibition of human immunodeficiency virus type 1 (HIV-1) nuclear import via Vpr-Imporin alpha interactions as a novel HIV-1 therapy, *Biochem. Biophys. Res. Commun.* 380 (2009) 838–843.
- [11] S. Bourbigot, H. Beltz, J. Denis, N. Morellet, B.P. Roques, Y. Mely, S. Bouaziz, The C-terminal domain of the HIV-1 regulatory protein Vpr adopts an antiparallel dimeric structure in solution via its leucine-zipper-like domain, *Biochem. J.* 387 (2005) 333–341.
- [12] P. Henklein, K. Bruns, M.P. Sherman, U. Tessmer, K. Licha, J. Kopp, C.M. de Noronha, W.C. Greene, V. Wray, U. Schubert, Functional and structural characterization of synthetic HIV-1 Vpr that transduces cells, localizes to the nucleus, and induces G2 cell cycle arrest, *J. Biol. Chem.* 275 (2000) 32016–32026.
- [13] W. Schuler, K. Wecker, H. de Rocquigny, Y. Baudat, J. Sire, B.P. Roques, NMR structure of the (52–96) C-terminal domain of the HIV-1 regulatory protein Vpr: molecular insights into its biological functions, *J. Mol. Biol.* 285 (1999) 2105–2117.
- [14] K. Wecker, B.P. Roques, NMR structure of the (1–51) N-terminal domain of the HIV-1 regulatory protein Vpr, *Eur. J. Biochem.* 266 (1999) 359–369.
- [15] N. Morellet, S. Bouaziz, P. Petitjean, B.P. Roques, NMR structure of the HIV-1 regulatory protein VPR, *J. Mol. Biol.* 327 (2003) 215–227.
- [16] N. Morellet, B.P. Roques, S. Bouaziz, Structure-function relationship of Vpr: biological implications, *Curr. HIV Res.* 7 (2009) 184–210.
- [17] G. Jacquot, E. Le Rouzic, A. David, J. Mazzolini, J. Bouchet, S. Bouaziz, F. Niedergang, G. Pancino, S. Benichou, Localization of HIV-1 Vpr to the nuclear envelope: Impact on Vpr functions and virus replication in macrophages, *Retrovirology* 4 (2007) 84.
- [18] D.L. Bolton, M.J. Lenardo, Vpr cytopathicity independent of G2/M cell cycle arrest in human immunodeficiency virus type 1-infected CD4+ T cells, *J. Virol.* 81 (2007) 8878–8890.
- [19] J.V. Fritz, P. Didier, J.P. Clamme, E. Schaub, D. Muriaux, C. Cabanne, N. Morellet, S. Bouaziz, J.L. Darlix, Y. Mely, H. de Rocquigny, Direct Vpr-Vpr interaction in cells monitored by two photon fluorescence correlation spectroscopy and fluorescence lifetime imaging, *Retrovirology* 5 (2008) 87.
- [20] J.V. Fritz, D. Dujardin, J. Godet, P. Didier, J. De Mey, J.L. Darlix, Y. Mely, H. de Rocquigny, HIV-1 Vpr oligomerization but not that of Gag directs the interaction between Vpr and Gag, *J. Virol.* 84 (2010) 1585–1596.
- [21] J.V. Fritz, L. Briant, Y. Mely, S. Bouaziz, H. de Rocquigny, HIV-1 Viral Protein R: from structure to function, *Future Virol.* 5 (2010) 607–625.
- [22] C. Arunagiri, I. Macreadie, D. Hewish, A. Azad, A C-terminal domain of HIV-1 accessory protein Vpr is involved in penetration, mitochondrial dysfunction and apoptosis of human CD4+ lymphocytes, *Apoptosis* 2 (1997) 69–76.
- [23] E.N. Sabbah, B.P. Roques, Critical implication of the (70–96) domain of human immunodeficiency virus type 1 Vpr protein in apoptosis of primary rat cortical and striatal neurons, *J. Neurovirology* 11 (2005) 489–502.
- [24] A.N. Godet, J. Guernon, A. Croset, X. Cayla, P.B. Falanga, J.H. Colle, A. Garcia, PP2A1 binding, cell transducing and apoptotic properties of Vpr(77–92): a new functional domain of HIV-1 Vpr proteins, *PLoS One* 5 (2010) e13760.
- [25] M.P. Sherman, U. Schubert, S.A. Williams, C.M. de Noronha, J.F. Kreisberg, P. Henklein, W.C. Greene, HIV-1 Vpr displays natural protein-transducing properties: implications for viral pathogenesis, *Virology* 302 (2002) 95–105.
- [26] K. Siddiqui, L. Del Valle, N. Morellet, J. Cui, M. Ghafouri, R. Mukerjee, K. Urbanska, S. Fan, C.B. Pattillo, S.L. Deshmane, M.F. Kiani, R. Ansari, K. Khalili, B.P. Roques, K. Reiss, S. Bouaziz, S. Amiri, A. Srinivasan, B.E. Sawaya, Molecular mimicry in inducing DNA damage between HIV-1 Vpr and the anticancer agent, cisplatin, *Oncogene* 27 (2008) 32–43.
- [27] K. Muthumani, V.M. Lambert, M. Shanmugam, K.P. Thieu, A.Y. Choo, J.C. Chung, A. Satishchandran, J.J. Kim, D.B. Weiner, K.E. Ugen, Anti-tumor activity mediated by protein and peptide transduction of HIV viral protein R (Vpr), *Cancer Biol. Therapy* 8 (2009) 180–187.
- [28] M. Cadene, C. Boudier, G.D. de Marcillac, J.G. Bieth, Influence of low molecular mass heparin on the kinetics of neutrophil elastase inhibition by mucus proteinase inhibitor, *J. Biol. Chem.* 270 (1995) 13204–13209.
- [29] H. de Rocquigny, A. Caneparo, T. Delaunay, J. Bischerour, J.F. Mouscadet, B.P. Roques, Interactions of the C-terminus of viral protein R with nucleic acids are modulated by its N-terminus, *Eur. J. Biochem.* 267 (2000) 3654–3660.
- [30] R. Zidovetzki, I. Levitan, Use of cyclodextrins to manipulate plasma membrane cholesterol content: evidence, misconceptions and control strategies, *Biochim. Biophys. Acta* 1768 (2007) 1311–1324.
- [31] F. Gabor, U. Klaussegger, M. Wirth, The interaction between wheat germ agglutinin and other plant lectins with prostate cancer cells Du-145, *Int. J. Pharm.* 221 (2001) 35–47.
- [32] J. Azoulay, J.P. Clamme, J.L. Darlix, B.P. Roques, Y. Mely, Destabilization of the HIV-1 complementary sequence of TAR by the nucleocapsid protein through activation of conformational fluctuations, *J. Mol. Biol.* 326 (2003) 691–700.
- [33] J.P. Clamme, J. Azoulay, Y. Mely, Monitoring of the formation and dissociation of polyethylenimine/DNA complexes by two photon fluorescence correlation spectroscopy, *Biophys. J.* 84 (2003) 1960–1968.
- [34] T.D. Madden, M.B. Bally, M.J. Hope, P.R. Cullis, H.P. Schieren, A.S. Janoff, Protection of large unilamellar vesicles by trehalose during dehydration: retention of vesicle contents, *Biochim. Biophys. Acta* 817 (1985) 67–74.
- [35] S. Bernacchi, Y. Mely, Exciton interaction in molecular beacons: a sensitive sensor for short range modifications of the nucleic acid structure, *Nucleic Acids Res.* 29 (2001) E62–62.
- [36] M. Kasha, Energy transfer, charge transfer, and proton transfer in molecular composite systems, *Basic Life Sci.* 58 (1991) 231–251 (discussion 251–235).
- [37] B.Z. Packard, D.D. Toptygin, A. Komoriya, L. Brand, Profluorescent protease substrates: intramolecular dimers described by the exciton model, *Proc. Natl. Acad. Sci. U.S.A.* 93 (1996) 11640–11645.
- [38] J.P. Richard, K. Melikov, E. Vives, C. Ramos, B. Verbeure, M.J. Gait, L.V. Chernomordik, B. Lebleu, Cell-penetrating peptides. A reevaluation of the mechanism of cellular uptake, *J. Biol. Chem.* 278 (2003) 585–590.
- [39] F. Duchardt, M. Fotin-Mlczek, H. Schwarz, R. Fischer, R. Brock, A comprehensive model for the cellular uptake of cationic cell-penetrating peptides, *Traffic (Copenhagen, Denmark)* 8 (2007) 848–866.
- [40] L.N. Patel, J.L. Zaro, W.C. Shen, Cell penetrating peptides: intracellular pathways and pharmaceutical perspectives, *Pharm. Res.* 24 (2007) 1977–1992.
- [41] M. Rusnati, D. Coltrini, P. Oreste, G. Zoppetti, A. Albini, D. Noonan, F. D'Adda di Fagagna, M. Giacca, M. Presta, Interaction of HIV-1 Tat protein with heparin. Role of the backbone structure, sulfation, and size, *J. Biol. Chem.* 272 (1997) 11313–11320.
- [42] S.L. Schmid, Clathrin-coated vesicle formation and protein sorting: an integrated process, *Annu. Rev. Biochem.* 66 (1997) 511–548.
- [43] R.G. Anderson, The caveolae membrane system, *Annu. Rev. Biochem.* 67 (1998) 199–225.
- [44] A. Fittipaldi, A. Ferrari, M. Zoppe, C. Arcangeli, V. Pellegrini, F. Beltram, M. Giacca, Cell membrane lipid rafts mediate caveolar endocytosis of HIV-1 Tat fusion proteins, *J. Biol. Chem.* 278 (2003) 34141–34149.
- [45] E. Coeytaux, D. Coulaud, E. Le Cam, O. Danos, A. Kichler, The cationic amphiphilic alpha-helix of HIV-1 viral protein R (Vpr) binds to nucleic acids, permeabilizes membranes, and efficiently transfects cells, *J. Biol. Chem.* 278 (2003) 18110–18116.
- [46] J.A. Swanson, C. Watts, Macropinocytosis, *Trends Cell Biol.* 5 (1995) 424–428.
- [47] N.Q. Liu, A.S. Lossinsky, W. Popik, X. Li, C. Gujuluva, B. Kriederman, J. Roberts, T. Pushkarsky, M. Bukrinsky, M. Witte, M. Weinand, M. Fiala, Human immunodeficiency virus type 1 enters brain microvascular endothelium by macropinocytosis dependent on lipid rafts and the mitogen-activated protein kinase signaling pathway, *J. Virol.* 76 (2002) 6689–6700.
- [48] N. Schmidt, A. Mishra, G.H. Lai, G.C. Wong, Arginine-rich cell-penetrating peptides, *FEBS Lett.* 584 (2010) 1806–1813.
- [49] Z.V. Leonenko, E. Finot, H. Ma, T.E. Dahms, D.T. Cramb, Investigation of temperature-induced phase transitions in DOPC and DPPC phospholipid bilayers using temperature-controlled scanning force microscopy, *Biophys. J.* 86 (2004) 3783–3793.
- [50] I. Vergne, P. Constant, G. Laneelle, Phagosomal pH determination by dual fluorescence flow cytometry, *Anal. Biochem.* 255 (1998) 127–132.
- [51] A. Dinter, E.G. Berger, Golgi-disturbing agents, *Histochem. Cell Biol.* 109 (1998) 571–590.
- [52] S. Aubry, B. Aussedat, D. Delaroche, C.Y. Jiao, G. Bolbach, S. Lavielle, G. Chassaing, S. Sagan, F. Burlina, MALDI-TOF mass spectrometry: a powerful tool to study the internalization of cell-penetrating peptides, *Biochim. Biophys. Acta* (2009).
- [53] G. Krishnamoorthy, G. Dupontail, Y. Mely, Structure and dynamics of condensed DNA probed by 1,1'-(4,4,8,8-tetramethyl-4,8-diazaundecamethylene)bis[4-[[3-methylbenzyl-3-oxazol-2-yl]methylidene]-1,4-dihydroquinolinium] tetraiodide fluorescence, *Biochemistry* 41 (2002) 15277–15287.
- [54] J.R. Lakowicz, Fluorescence spectroscopic investigations of the dynamic properties of proteins, membranes and nucleic acids, *J. Biochem. Biophys. Methods* 2 (1980) 91–119.
- [55] R.Y. Zhao, D. Liang, G. Li, C.W. Larrimore, B.L. Mirkin, Anti-cancer effect of HIV-1 viral protein R on doxorubicin resistant neuroblastoma, *PLoS One* 5 (2010) e11466.



- [56] T. Taguchi, M. Shimura, Y. Osawa, Y. Suzuki, I. Mizoguchi, K. Niino, F. Takaku, Y. Ishizaka, Nuclear trafficking of macromolecules by an oligopeptide derived from Vpr of human immunodeficiency virus type-1, *Biochem. Biophys. Res. Commun.* 320 (2004) 18–26.
- [57] V. Tiriveedhi, P. Butko, A fluorescence spectroscopy study on the interactions of the TAT-PTD peptide with model lipid membranes, *Biochemistry* 46 (2007) 3888–3895.
- [58] C.M. de Noronha, M.P. Sherman, H.W. Lin, M.V. Cavois, R.D. Moir, R.D. Goldman, W.C. Greene, Dynamic disruptions in nuclear envelope architecture and integrity induced by HIV-1 Vpr, *Science* 294 (2001) 1105–1108.
- [59] S.C. Piller, P. Jans, P.W. Gage, D.A. Jans, Extracellular HIV-1 virus protein R causes a large inward current and cell death in cultured hippocampal neurons: implications for AIDS pathology, *Proc. Natl. Acad. Sci. U.S.A.* 95 (1998) 4595–4600.
- [60] A. Ziegler, Thermodynamic studies and binding mechanisms of cell-penetrating peptides with lipids and glycosaminoglycans, *Advan. Drug Delivery Rev.* 60 (2008) 580–597.
- [61] S.C. Piller, G.D. Ewart, A. Premkumar, G.B. Cox, P.W. Gage, Vpr protein of human immunodeficiency virus type 1 forms cation-selective channels in planar lipid bilayers, *Proc. Natl. Acad. Sci. U.S.A.* 93 (1996) 111–115.
- [62] S. Sandgren, F. Cheng, M. Belting, Nuclear targeting of macromolecular polyanions by an HIV-Tat derived peptide. Role for cell-surface proteoglycans, *J. Biol. Chem.* 277 (2002) 38877–38883.
- [63] E. Goncalves, E. Kitas, J. Seelig, Binding of oligoarginine to membrane lipids and heparan sulfate: structural and thermodynamic characterization of a cell-penetrating peptide, *Biochemistry* 44 (2005) 2692–2702.
- [64] A. Ziegler, J. Seelig, Interaction of the protein transduction domain of HIV-1 TAT with heparan sulfate: binding mechanism and thermodynamic parameters, *Biophys. J.* 86 (2004) 254–263.
- [65] A. Ziegler, X.L. Blatter, A. Seelig, J. Seelig, Protein transduction domains of HIV-1 and SIV TAT interact with charged lipid vesicles. Binding mechanism and thermodynamic analysis, *Biochemistry* 42 (2003) 9185–9194.
- [66] R. Fischer, K. Kohler, M. Fotin-Mlecsek, R. Brock, A stepwise dissection of the intracellular fate of cationic cell-penetrating peptides, *J. Biol. Chem.* 279 (2004) 12625–12635.
- [67] A. Vendeville, F. Rayne, A. Bonhoure, N. Bettache, P. Montcourrier, B. Beaumelle, HIV-1 Tat enters T cells using coated pits before translocating from acidified endosomes and eliciting biological responses, *Mol. Biol. Cell* 15 (2004) 2347–2360.
- [68] T.B. Potocky, A.K. Menon, S.H. Gellman, Cytoplasmic and nuclear delivery of a TAT-derived peptide and a beta-peptide after endocytic uptake into HeLa cells, *J. Biol. Chem.* 278 (2003) 50188–50194.
- [69] R. Trehin, H.M. Nielsen, H.G. Jahnke, U. Krauss, A.G. Beck-Sickinger, H.P. Merkle, Metabolic cleavage of cell-penetrating peptides in contact with epithelial models: human calcitonin (hCT)-derived peptides, Tat(47-57) and penetratin(43-58), *Biochem. J.* 382 (2004) 945–956.
- [70] H. Binder, G. Lindblom, Charge-dependent translocation of the Trojan peptide penetratin across lipid membranes, *Biophys. J.* 85 (2003) 982–995.
- [71] D. Persson, P.E. Thoren, M. Herner, P. Lincoln, B. Norden, Application of a novel analysis to measure the binding of the membrane-translocating peptide penetratin to negatively charged liposomes, *Biochemistry* 42 (2003) 421–429.
- [72] A. Lamaziere, C. Wolf, O. Lambert, G. Chassaing, G. Trugnan, J. Ayala-Sanmartin, The homeodomain derived peptide Penetratin induces curvature of fluid membrane domains, *PLoS One* 3 (2008) e1938.
- [73] D. Terrone, S.L. Sang, L. Roudaia, J.R. Silvius, Penetratin and related cell-penetrating cationic peptides can translocate across lipid bilayers in the presence of a transbilayer potential, *Biochemistry* 42 (2003) 13787–13799.
- [74] H.D. Herce, A.E. Garcia, Molecular dynamics simulations suggest a mechanism for translocation of the HIV-1 TAT peptide across lipid membranes, *Proc. Natl. Acad. Sci. U.S.A.* 104 (2007) 20805–20810.
- [75] R. Rothe, L. Liguori, A. Villegas-Mendez, B. Marques, D. Grunwald, E. Drouet, J.L. Lenormand, Characterization of the cell-penetrating properties of the Epstein-Barr virus ZEBRA trans-activator, *J. Biol. Chem.* 285 (2010) 20224–20233.

## VII- REFERENCES BIBLIOGRAPHIQUES

- Abbady AQ, Bronner C, Bathami K, Muller CD, Jeanblanc M, Mathieu E, Klein JP, Candolfi E and Mousli M (2005) TCR pathway involves ICBP90 gene down-regulation via E2F binding sites. *Biochemical pharmacology* 70:570-579.
- Abbady AQ, Bronner C, Trotzier MA, Hopfner R, Bathami K, Muller CD, Jeanblanc M and Mousli M (2003) ICBP90 expression is downregulated in apoptosis-induced Jurkat cells. *Annals of the New York Academy of Sciences* 1010:300-303.
- Abbott DW, Chadwick BP, Thambirajah AA and Ausio J (2005) Beyond the Xi: macroH2A chromatin distribution and post-translational modification in an avian system. *The Journal of biological chemistry* 280:16437-16445.
- Abusnina A, Alhosin M, Keravis T, Muller CD, Fuhrmann G, Bronner C and Lugnier C (2011) Down-regulation of cyclic nucleotide phosphodiesterase PDE1A is the key event of p73 and UHRF1 deregulation in thymoquinone-induced acute lymphoblastic leukemia cell apoptosis. *Cellular signalling* 23:152-160.
- Achour M, Fuhrmann G, Alhosin M, Ronde P, Chataigneau T, Mousli M, Schini-Kerth VB and Bronner C (2009) UHRF1 recruits the histone acetyltransferase Tip60 and controls its expression and activity. *Biochemical and biophysical research communications* 390:523-528.
- Achour M, Jacq X, Ronde P, Alhosin M, Charlot C, Chataigneau T, Jeanblanc M, Macaluso M, Giordano A, Hughes AD, Schini-Kerth VB and Bronner C (2008) The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of VEGF gene expression. *Oncogene* 27:2187-2197.
- Alhosin M, Abusnina A, Achour M, Sharif T, Muller C, Peluso J, Chataigneau T, Lugnier C, Schini-Kerth VB, Bronner C and Fuhrmann G (2010) Induction of apoptosis by thymoquinone in lymphoblastic leukemia Jurkat cells is mediated by a p73-dependent pathway which targets the epigenetic integrator UHRF1. *Biochemical pharmacology* 79:1251-1260.
- Ambros V (2004) The functions of animal microRNAs. *Nature* 431:350-355.
- Antequera F (2003) Structure, function and evolution of CpG island promoters. *Cell Mol Life Sci* 60:1647-1658.

- Aoki A, Suetake I, Miyagawa J, Fujio T, Chijiwa T, Sasaki H and Tajima S (2001) Enzymatic properties of de novo-type mouse DNA (cytosine-5) methyltransferases. *Nucleic acids research* 29:3506-3512.
- Arand J, Spieler D, Karius T, Branco MR, Meilinger D, Meissner A, Jenuwein T, Xu G, Leonhardt H, Wolf V and Walter J (2012) In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases. *PLoS genetics* 8:e1002750.
- Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T and Hannon GJ (2008) A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Molecular cell* 31:785-799.
- Arima Y, Hirota T, Bronner C, Mousli M, Fujiwara T, Niwa S, Ishikawa H and Saya H (2004) Down-regulation of nuclear protein ICBP90 by p53/p21Cip1/WAF1-dependent DNA-damage checkpoint signals contributes to cell cycle arrest at G1/S transition. *Genes Cells* 9:131-142.
- Arita K, Ariyoshi M, Tochio H, Nakamura Y and Shirakawa M (2008) Recognition of hemimethylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature* 455:818-821.
- Arita K, Isogai S, Oda T, Unoki M, Sugita K, Sekiyama N, Kuwata K, Hamamoto R, Tochio H, Sato M, Ariyoshi M and Shirakawa M (2012) Recognition of modification status on a histone H3 tail by linked histone reader modules of the epigenetic regulator UHRF1. *Proceedings of the National Academy of Sciences of the United States of America* 109:12950-12955.
- Arrowsmith CH, Bountra C, Fish PV, Lee K and Schapira M (2012) Epigenetic protein families: a new frontier for drug discovery. *Nat Rev Drug Discov* 11:384-400.
- Avvakumov GV, Walker JR, Xue S, Li Y, Duan S, Bronner C, Arrowsmith CH and Dhe-Paganon S (2008) Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. *Nature* 455:822-825.
- Babbio F, Pistore C, Curti L, Castiglioni I, Kunderfranco P, Brino L, Oudet P, Seiler R, Thalman GN, Roggero E, Sarti M, Pinton S, Mello-Grand M, Chiorino G, Catapano CV, Carbone GM and Bonapace IM (2012) The SRA protein UHRF1 promotes epigenetic crosstalks and is involved in prostate cancer progression. *Oncogene*.
- Baccarelli A, Wright RO, Bollati V, Tarantini L, Litonjua AA, Suh HH, Zanobetti A, Sparrow D, Vokonas PS and Schwartz J (2009) Rapid DNA methylation changes after

- exposure to traffic particles. *American journal of respiratory and critical care medicine* 179:572-578.
- Badea MG and Brand L (1979) Time-resolved fluorescence measurements. *Methods in enzymology* 61:378-425.
- Ballestar E and Esteller M (2008) SnapShot: the human DNA methylome in health and disease. *Cell* 135:1144-1144 e1141.
- Bannister AJ and Kouzarides T (2011) Regulation of chromatin by histone modifications. *Cell research* 21:381-395.
- Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M and Kouzarides T (2010) Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* 143:470-484.
- Basu R and Zhang LF (2011) X chromosome inactivation: a silence that needs to be broken. *Genesis* 49:821-834.
- Bedford MT and Clarke SG (2009) Protein arginine methylation in mammals: who, what, and why. *Molecular cell* 33:1-13.
- Bell AC and Felsenfeld G (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* 405:482-485.
- Berger SL, Kouzarides T, Shiekhatar R and Shilatifard A (2009) An operational definition of epigenetics. *Genes & development* 23:781-783.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL and Lander ES (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125:315-326.
- Bernstein E and Allis CD (2005) RNA meets chromatin. *Genes & development* 19:1635-1655.
- Bestor TH (2000) The DNA methyltransferases of mammals. *Human molecular genetics* 9:2395-2402.
- Bhutani N, Burns DM and Blau HM (2011) DNA demethylation dynamics. *Cell* 146:866-872.
- Bianchi C and Zangi R (2012) How to Distinguish methyl-Cytosine from Cytosine with High Fidelity. *Journal of molecular biology*.
- Binda O, LeRoy G, Bua DJ, Garcia BA, Gozani O and Richard S (2010) Trimethylation of histone H3 lysine 4 impairs methylation of histone H3 lysine 9: regulation of lysine



- methyltransferases by physical interaction with their substrates. *Epigenetics* 5:767-775.
- Bird A (2002) DNA methylation patterns and epigenetic memory. *Genes & development* 16:6-21.
- Bloomfield VA, Crothers DM and Tinoco I, Jr. (2000) *Nucleic acids: structures, properties, and functions*, University Science Books, Sausalito, California.
- Bollati V, Baccarelli A, Hou L, Bonzini M, Fustinoni S, Cavallo D, Byun HM, Jiang J, Marinelli B, Pesatori AC, Bertazzi PA and Yang AS (2007) Changes in DNA methylation patterns in subjects exposed to low-dose benzene. *Cancer research* 67:876-880.
- Bombarda E, Ababou A, Vuilleumier C, Gerard D, Roques BP, Piemont E and Mely Y (1999) Time-resolved fluorescence investigation of the human immunodeficiency virus type 1 nucleocapsid protein: influence of the binding of nucleic acids. *Biophysical journal* 76:1561-1570.
- Borgel J, Guibert S, Li Y, Chiba H, Schubeler D, Sasaki H, Forne T and Weber M (2010) Targets and dynamics of promoter DNA methylation during early mouse development. *Nature genetics* 42:1093-1100.
- Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S and Jacobsen SE (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science (New York, NY)* 317:1760-1764.
- Bourc'his D and Bestor TH (2004) Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* 431:96-99.
- Bourc'his D, Xu GL, Lin CS, Bollman B and Bestor TH (2001) Dnmt3L and the establishment of maternal genomic imprints. *Science (New York, NY)* 294:2536-2539.
- Braconi C, Huang N and Patel T (2010) MicroRNA-dependent regulation of DNA methyltransferase-1 and tumor suppressor gene expression by interleukin-6 in human malignant cholangiocytes. *Hepatology (Baltimore, Md)* 51:881-890.
- Branco MR, Ficiz G and Reik W (2011) Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat Rev Genet* 13:7-13.
- Bronner C (2011) Control of DNMT1 abundance in epigenetic inheritance by acetylation, ubiquitylation, and the histone code. *Science signaling* 4:pe3.

- Bronner C, Achour M, Arima Y, Chataigneau T, Saya H and Schini-Kerth VB (2007a) The UHRF family: oncogenes that are drugable targets for cancer therapy in the near future? *Pharmacology & therapeutics* 115:419-434.
- Bronner C, Chataigneau T, Schini-Kerth VB and Landry Y (2007b) The "Epigenetic Code Replication Machinery", ECREM: a promising drugable target of the epigenetic cell memory. *Current medicinal chemistry* 14:2629-2641.
- Bronner C, Fuhrmann G, Chedin FL, Macaluso M and Dhe-Paganon S (2010) UHRF1 Links the Histone code and DNA Methylation to ensure Faithful Epigenetic Memory Inheritance. *Genetics & epigenetics* 2009:29-36.
- Cairns BR (2009) The logic of chromatin architecture and remodelling at promoters. *Nature* 461:193-198.
- Calvanese V, Fernandez AF, Urdinguio RG, Suarez-Alvarez B, Mangas C, Perez-Garcia V, Bueno C, Montes R, Ramos-Mejia V, Martinez-Cambor P, Ferrero C, Assenov Y, Bock C, Menendez P, Carrera AC, Lopez-Larrea C and Fraga MF (2012) A promoter DNA demethylation landscape of human hematopoietic differentiation. *Nucleic acids research* 40:116-131.
- Cao Q, Mani RS, Ateeq B, Dhanasekaran SM, Asangani IA, Prensner JR, Kim JH, Brenner JC, Jing X, Cao X, Wang R, Li Y, Dahiya A, Wang L, Pandhi M, Lonigro RJ, Wu YM, Tomlins SA, Palanisamy N, Qin Z, Yu J, Maher CA, Varambally S and Chinnaiyan AM (2011) Coordinated regulation of polycomb group complexes through microRNAs in cancer. *Cancer cell* 20:187-199.
- Cedar H and Bergman Y (2009) Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet* 10:295-304.
- Chedin F, Lieber MR and Hsieh CL (2002) The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proceedings of the National Academy of Sciences of the United States of America* 99:16916-16921.
- Chen T, Ueda Y, Dodge JE, Wang Z and Li E (2003) Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Molecular and cellular biology* 23:5594-5605.
- Chen ZX, Mann JR, Hsieh CL, Riggs AD and Chedin F (2005) Physical and functional interactions between the human DNMT3L protein and members of the de novo methyltransferase family. *Journal of cellular biochemistry* 95:902-917.

- Chen ZX and Riggs AD (2011) DNA methylation and demethylation in mammals. *The Journal of biological chemistry* 286:18347-18353.
- Cheng X (1995) Structure and function of DNA methyltransferases. *Annual review of biophysics and biomolecular structure* 24:293-318.
- Cheng X and Roberts RJ (2001) AdoMet-dependent methylation, DNA methyltransferases and base flipping. *Nucleic acids research* 29:3784-3795.
- Chodavarapu RK, Feng S, Bernatavichute YV, Chen PY, Stroud H, Yu Y, Hetzel JA, Kuo F, Kim J, Cokus SJ, Casero D, Bernal M, Huijser P, Clark AT, Kramer U, Merchant SS, Zhang X, Jacobsen SE and Pellegrini M (2010) Relationship between nucleosome positioning and DNA methylation. *Nature* 466:388-392.
- Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL, Nelson HH, Karagas MR, Padbury JF, Bueno R, Sugarbaker DJ, Yeh RF, Wiencke JK and Kelsey KT (2009) Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. *PLoS genetics* 5:e1000602.
- Chuang LS, Ian HI, Koh TW, Ng HH, Xu G and Li BF (1997) Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science (New York, NY)* 277:1996-2000.
- Citterio E, Papait R, Nicassio F, Vecchi M, Gomiero P, Mantovani R, Di Fiore PP and Bonapace IM (2004) Np95 is a histone-binding protein endowed with ubiquitin ligase activity. *Molecular and cellular biology* 24:2526-2535.
- Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D, Abramowitz LK, Bartolomei MS, Rambow F, Bassi MR, Bruno T, Fanciulli M, Renner C, Klein-Szanto AJ, Matsumoto Y, Kobi D, Davidson I, Alberti C, Larue L and Bellacosa A (2011) Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* 146:67-79.
- Costanzi C and Pehrson JR (2001) MACROH2A2, a new member of the MARCOH2A core histone family. *The Journal of biological chemistry* 276:21776-21784.
- Costello JF, Fruhwald MC, Smiraglia DJ, Rush LJ, Robertson GP, Gao X, Wright FA, Feramisco JD, Peltomaki P, Lang JC, Schuller DE, Yu L, Bloomfield CD, Caligiuri MA, Yates A, Nishikawa R, Su Huang H, Petrelli NJ, Zhang X, O'Dorisio MS, Held WA, Cavenee WK and Plass C (2000) Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nature genetics* 24:132-138.

- Creyghton MP, Markoulaki S, Levine SS, Hanna J, Lodato MA, Sha K, Young RA, Jaenisch R and Boyer LA (2008) H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment. *Cell* 135:649-661.
- Crnogorac-Jurcevic T, Gangeswaran R, Bhakta V, Capurso G, Lattimore S, Akada M, Sunamura M, Prime W, Campbell F, Brentnall TA, Costello E, Neoptolemos J and Lemoine NR (2005) Proteomic analysis of chronic pancreatitis and pancreatic adenocarcinoma. *Gastroenterology* 129:1454-1463.
- Daskalos A, Oleksiewicz U, Folia A, Nikolaidis G, Xinarianos G, Gosney JR, Malliri A, Field JK and Liloglou T (2011) UHRF1-mediated tumor suppressor gene inactivation in nonsmall cell lung cancer. *Cancer* 117:1027-1037.
- Davis TL, Yang GJ, McCarrey JR and Bartolomei MS (2000) The H19 methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. *Human molecular genetics* 9:2885-2894.
- Dawson MA and Kouzarides T (2012) Cancer epigenetics: from mechanism to therapy. *Cell* 150:12-27.
- Daxinger L and Whitelaw E (2012) Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nat Rev Genet* 13:153-162.
- Delagoutte B, Lallous N, Birck C, Oudet P and Samama JP (2008) Expression, purification, crystallization and preliminary crystallographic study of the SRA domain of the human UHRF1 protein. *Acta crystallographica* 64:922-925.
- Denis H, Ndlovu MN and Fuks F (2011) Regulation of mammalian DNA methyltransferases: a route to new mechanisms. *EMBO reports* 12:647-656.
- Dhayalan A, Rajavelu A, Rathert P, Tamas R, Jurkowska RZ, Ragozin S and Jeltsch A (2010) The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *The Journal of biological chemistry* 285:26114-26120.
- Didier P, Sharma KK and Mély Y (2011) Fluorescence Techniques to Characterise Ligand Binding to Proteins in *Biophysical Approaches Determining Ligand Binding to Biomolecular Targets: Detection, Measurement and Modelling* (Podjarny A, Dejaegere AP and Kieffer B eds) pp 156-199, The Royal Society of Chemistry.
- Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB, Askarian-Amiri ME, Ru K, Solda G, Simons C, Sunkin SM, Crowe ML, Grimmond SM, Perkins AC and Mattick JS (2008) Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome research* 18:1433-1445.

- Dodge JE, Okano M, Dick F, Tsujimoto N, Chen T, Wang S, Ueda Y, Dyson N and Li E (2005) Inactivation of Dnmt3b in mouse embryonic fibroblasts results in DNA hypomethylation, chromosomal instability, and spontaneous immortalization. *The Journal of biological chemistry* 280:17986-17991.
- Dong KB, Maksakova IA, Mohn F, Leung D, Appanah R, Lee S, Yang HW, Lam LL, Mager DL, Schubeler D, Tachibana M, Shinkai Y and Lorincz MC (2008) DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *The EMBO journal* 27:2691-2701.
- Dorigo B, Schalch T, Kulangara A, Duda S, Schroeder RR and Richmond TJ (2004) Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science (New York, NY)* 306:1571-1573.
- Du Z, Song J, Wang Y, Zhao Y, Guda K, Yang S, Kao HY, Xu Y, Willis J, Markowitz SD, Sedwick D, Ewing RM and Wang Z (2010) DNMT1 stability is regulated by proteins coordinating deubiquitination and acetylation-driven ubiquitination. *Science signaling* 3:ra80.
- Duan Q, Chen H, Costa M and Dai W (2008) Phosphorylation of H3S10 blocks the access of H3K9 by specific antibodies and histone methyltransferase. Implication in regulating chromatin dynamics and epigenetic inheritance during mitosis. *The Journal of biological chemistry* 283:33585-33590.
- Duursma AM, Kedde M, Schrier M, le Sage C and Agami R (2008) miR-148 targets human DNMT3b protein coding region. *RNA (New York, NY)* 14:872-877.
- Easwaran HP, Schermelleh L, Leonhardt H and Cardoso MC (2004) Replication-independent chromatin loading of Dnmt1 during G2 and M phases. *EMBO reports* 5:1181-1186.
- Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA, Haefliger C, Horton R, Howe K, Jackson DK, Kunde J, Koenig C, Liddle J, Niblett D, Otto T, Pettett R, Seemann S, Thompson C, West T, Rogers J, Olek A, Berlin K and Beck S (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. *Nature genetics* 38:1378-1385.
- Edwards CA and Ferguson-Smith AC (2007) Mechanisms regulating imprinted genes in clusters. *Current opinion in cell biology* 19:281-289.
- Egger G, Jeong S, Escobar SG, Cortez CC, Li TW, Saito Y, Yoo CB, Jones PA and Liang G (2006) Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic

- knockouts suggests an essential role for DNMT1 in cell survival. *Proceedings of the National Academy of Sciences of the United States of America* 103:14080-14085.
- Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA and Gehrke C (1982) Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic acids research* 10:2709-2721.
- Eisinger J and Navon G (1969) Fluorescence quenching and isotope effect of tryptophan. *J Chem Phys* 50:2069-2077.
- El-Maarri O, Buiting K, Peery EG, Kroisel PM, Balaban B, Wagner K, Urman B, Heyd J, Lich C, Brannan CI, Walter J and Horsthemke B (2001) Maternal methylation imprints on human chromosome 15 are established during or after fertilization. *Nature genetics* 27:341-344.
- Epsztejn-Litman S, Feldman N, Abu-Remaileh M, Shufaro Y, Gerson A, Ueda J, Deplus R, Fuks F, Shinkai Y, Cedar H and Bergman Y (2008) De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nature structural & molecular biology* 15:1176-1183.
- Ernst J and Kellis M (2010) Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nature biotechnology* 28:817-825.
- Esteller M (2008) Epigenetics in cancer. *The New England journal of medicine* 358:1148-1159.
- Esteller M, Corn PG, Baylin SB and Herman JG (2001) A gene hypermethylation profile of human cancer. *Cancer research* 61:3225-3229.
- Esteve PO, Chin HG, Benner J, Feehery GR, Samaranyake M, Horwitz GA, Jacobsen SE and Pradhan S (2009) Regulation of DNMT1 stability through SET7-mediated lysine methylation in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* 106:5076-5081.
- Fabbri M, Garzon R, Cimmino A, Liu Z, Zanesi N, Callegari E, Liu S, Alder H, Costinean S, Fernandez-Cymering C, Volinia S, Guler G, Morrison CD, Chan KK, Marcucci G, Calin GA, Huebner K and Croce CM (2007) MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. *Proceedings of the National Academy of Sciences of the United States of America* 104:15805-15810.



- Falandry C, Fourel G, Galy V, Ristriani T, Horard B, Bensimon E, Salles G, Gilson E and Magdinier F (2010) CLLD8/KMT1F is a lysine methyltransferase that is important for chromosome segregation. *The Journal of biological chemistry* 285:20234-20241.
- Fan JY, Gordon F, Luger K, Hansen JC and Tremethick DJ (2002) The essential histone variant H2A.Z regulates the equilibrium between different chromatin conformational states. *Nature structural biology* 9:172-176.
- Fang F, Hodges E, Molaro A, Dean M, Hannon GJ and Smith AD (2012) Genomic landscape of human allele-specific DNA methylation. *Proceedings of the National Academy of Sciences of the United States of America* 109:7332-7337.
- Fang Z, Xing F, Bronner C, Teng Z and Guo Z (2009) ICBP90 mediates the ERK1/2 signaling to regulate the proliferation of Jurkat T cells. *Cellular immunology* 257:80-87.
- Fatemi M, Hermann A, Gowher H and Jeltsch A (2002) Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. *European journal of biochemistry / FEBS* 269:4981-4984.
- Fatemi M, Hermann A, Pradhan S and Jeltsch A (2001) The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. *Journal of molecular biology* 309:1189-1199.
- Federico M and Bagella L (2011) Histone deacetylase inhibitors in the treatment of hematological malignancies and solid tumors. *Journal of biomedicine & biotechnology* 2011:475641.
- Feil R and Fraga MF (2012) Epigenetics and the environment: emerging patterns and implications. *Nat Rev Genet* 13:97-109.
- Felle M, Joppien S, Nemeth A, Diermeier S, Thalhammer V, Dobner T, Kremmer E, Kappler R and Langst G (2011) The USP7/Dnmt1 complex stimulates the DNA methylation activity of Dnmt1 and regulates the stability of UHRF1. *Nucleic acids research* 39:8355-8365.
- Feltus FA, Lee EK, Costello JF, Plass C and Vertino PM (2003) Predicting aberrant CpG island methylation. *Proceedings of the National Academy of Sciences of the United States of America* 100:12253-12258.
- Feng S, Jacobsen SE and Reik W (2010) Epigenetic reprogramming in plant and animal development. *Science (New York, NY)* 330:622-627.

- Ferguson-Smith AC (2011) Genomic imprinting: the emergence of an epigenetic paradigm. *Nat Rev Genet* 12:565-575.
- Ferguson-Smith AC and Greally JM (2007) Epigenetics: perceptive enzymes. *Nature* 449:148-149.
- Fernandez-Capetillo O, Mahadevaiah SK, Celeste A, Romanienko PJ, Camerini-Otero RD, Bonner WM, Manova K, Burgoyne P and Nussenzweig A (2003) H2AX is required for chromatin remodeling and inactivation of sex chromosomes in male mouse meiosis. *Developmental cell* 4:497-508.
- Fernandez AF, Assenov Y, Martin-Subero JI, Balint B, Siebert R, Taniguchi H, Yamamoto H, Hidalgo M, Tan AC, Galm O, Ferrer I, Sanchez-Cespedes M, Villanueva A, Carmona J, Sanchez-Mut JV, Berdasco M, Moreno V, Capella G, Monk D, Ballestar E, Ropero S, Martinez R, Sanchez-Carbayo M, Prosper F, Agirre X, Fraga MF, Grana O, Perez-Jurado L, Mora J, Puig S, Prat J, Badimon L, Puca AA, Meltzer SJ, Lengauer T, Bridgewater J, Bock C and Esteller M (2012a) A DNA methylation fingerprint of 1628 human samples. *Genome research* 22:407-419.
- Fernandez AF, Huidobro C and Fraga MF (2012b) De novo DNA methyltransferases: oncogenes, tumor suppressors, or both? *Trends Genet* 28:474-479.
- Filion GJ, Zhenilo S, Salozhin S, Yamada D, Prokhortchouk E and Defossez PA (2006) A family of human zinc finger proteins that bind methylated DNA and repress transcription. *Molecular and cellular biology* 26:169-181.
- Fischle W, Tseng BS, Dormann HL, Ueberheide BM, Garcia BA, Shabanowitz J, Hunt DF, Funabiki H and Allis CD (2005) Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* 438:1116-1122.
- Förster T (1948) Zwischenmolekulare Energiewanderung und Fluoreszenz *Annalen der Physik* 437:55-75.
- Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K, Iyer NG, Perez-Rosado A, Calvo E, Lopez JA, Cano A, Calasanz MJ, Colomer D, Piris MA, Ahn N, Imhof A, Caldas C, Jenuwein T and Esteller M (2005) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nature genetics* 37:391-400.
- Frauer C, Hoffmann T, Bultmann S, Casa V, Cardoso MC, Antes I and Leonhardt H (2011) Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *PloS one* 6:e21306.

- Fuks F (2005) DNA methylation and histone modifications: teaming up to silence genes. *Current opinion in genetics & development* 15:490-495.
- Fuks F, Hurd PJ, Deplus R and Kouzarides T (2003a) The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic acids research* 31:2305-2312.
- Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP and Kouzarides T (2003b) The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *The Journal of biological chemistry* 278:4035-4040.
- Futscher BW, Oshiro MM, Wozniak RJ, Holtan N, Hanigan CL, Duan H and Domann FE (2002) Role for DNA methylation in the control of cell type specific maspin expression. *Nature genetics* 31:175-179.
- Garzon R, Liu S, Fabbri M, Liu Z, Heaphy CE, Callegari E, Schwind S, Pang J, Yu J, Muthusamy N, Havelange V, Volinia S, Blum W, Rush LJ, Perrotti D, Andreeff M, Bloomfield CD, Byrd JC, Chan K, Wu LC, Croce CM and Marcucci G (2009) MicroRNA-29b induces global DNA hypomethylation and tumor suppressor gene reexpression in acute myeloid leukemia by targeting directly DNMT3A and 3B and indirectly DNMT1. *Blood* 113:6411-6418.
- Ghoshal K, Datta J, Majumder S, Bai S, Kutay H, Motiwala T and Jacob ST (2005) 5-Aza-deoxycytidine induces selective degradation of DNA methyltransferase 1 by a proteasomal pathway that requires the KEN box, bromo-adjacent homology domain, and nuclear localization signal. *Molecular and cellular biology* 25:4727-4741.
- Glaser RL, Ramsay JP and Morison IM (2006) The imprinted gene and parent-of-origin effect database now includes parental origin of de novo mutations. *Nucleic acids research* 34:D29-31.
- Gluckman PD, Hanson MA, Buklijas T, Low FM and Beedle AS (2009) Epigenetic mechanisms that underpin metabolic and cardiovascular diseases. *Nat Rev Endocrinol* 5:401-408.
- Goll MG and Bestor TH (2005) Eukaryotic cytosine methyltransferases. *Annual review of biochemistry* 74:481-514.
- Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh CL, Zhang X, Golic KG, Jacobsen SE and Bestor TH (2006) Methylation of tRNA<sup>Asp</sup> by the DNA methyltransferase homolog Dnmt2. *Science (New York, NY)* 311:395-398.

- Gowher H and Jeltsch A (2001) Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites. *Journal of molecular biology* 309:1201-1208.
- Gowher H, Liebert K, Hermann A, Xu G and Jeltsch A (2005) Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. *The Journal of biological chemistry* 280:13341-13348.
- Goyal R, Reinhardt R and Jeltsch A (2006) Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase. *Nucleic acids research* 34:1182-1188.
- Gratton E and Limkeman M (1983) A continuously variable frequency cross-correlation phase fluorometer with picosecond resolution. *Biophysical journal* 44:315-324.
- Greiner VJ, Kovalenko L, Humbert N, Richert L, Birck C, Ruff M, Bronner C and Mely Y (en préparation) Site-specific characterization of the dynamics of UHRF1 SRA domain bound to 2-aminopurine labeled hemi-methylated oligonucleotides.
- Grigoryev SA, Arya G, Correll S, Woodcock CL and Schlick T (2009) Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. *Proceedings of the National Academy of Sciences of the United States of America* 106:13317-13322.
- Gronniger E, Weber B, Heil O, Peters N, Stab F, Wenck H, Korn B, Winnefeld M and Lyko F (2010) Aging and chronic sun exposure cause distinct epigenetic changes in human skin. *PLoS genetics* 6:e1000971.
- Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W, Xie ZG, Shi L, He X, Jin SG, Iqbal K, Shi YG, Deng Z, Szabo PE, Pfeifer GP, Li J and Xu GL (2011) The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* 477:606-610.
- Guan D, Factor D, Liu Y, Wang Z and Kao HY (2012) The epigenetic regulator UHRF1 promotes ubiquitination-mediated degradation of the tumor-suppressor protein promyelocytic leukemia protein. *Oncogene*.
- Guo JU, Su Y, Zhong C, Ming GL and Song H (2011) Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* 145:423-434.
- Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N, Bernstein BE, Kellis M, Regev A, Rinn JL and Lander ES (2009) Chromatin signature reveals

- over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 458:223-227.
- Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Cesari F, Lee C, Almouzni G, Schneider R and Surani MA (2008) Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* 452:877-881.
- Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W, Walter J and Surani MA (2002) Epigenetic reprogramming in mouse primordial germ cells. *Mechanisms of development* 117:15-23.
- Happel N and Doenecke D (2009) Histone H1 and its isoforms: contribution to chromatin structure and function. *Gene* 431:1-12.
- Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM and Tilghman SM (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 405:486-489.
- Hashimoto H, Horton JR, Zhang X, Bostick M, Jacobsen SE and Cheng X (2008) The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature* 455:826-829.
- Hashimoto H, Horton JR, Zhang X and Cheng X (2009) UHRF1, a modular multi-domain protein, regulates replication-coupled crosstalk between DNA methylation and histone modifications. *Epigenetics* 4:8-14.
- Hashimoto H, Vertino PM and Cheng X (2010) Molecular coupling of DNA methylation and histone methylation. *Epigenomics* 2:657-669.
- Hassa PO, Haenni SS, Elser M and Hottiger MO (2006) Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going? *Microbiol Mol Biol Rev* 70:789-829.
- Hata K, Okano M, Lei H and Li E (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development (Cambridge, England)* 129:1983-1993.
- He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song CX, Zhang K, He C and Xu GL (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science (New York, NY)* 333:1303-1307.
- Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, Slagboom PE and Lumey LH (2008) Persistent epigenetic differences associated with prenatal exposure to

- famine in humans. *Proceedings of the National Academy of Sciences of the United States of America* 105:17046-17049.
- Hendrich B and Tweedie S (2003) The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends Genet* 19:269-277.
- Hermann A, Goyal R and Jeltsch A (2004) The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *The Journal of biological chemistry* 279:48350-48359.
- Hervouet E, Lalier L, Debien E, Cheray M, Geairon A, Rogniaux H, Loussouarn D, Martin SA, Vallette FM and Cartron PF (2010a) Disruption of Dnmt1/PCNA/UHRF1 interactions promotes tumorigenesis from human and mice glial cells. *PloS one* 5:e11333.
- Hervouet E, Nadaradjane A, Gueguen M, Vallette FM and Cartron PF (2012) Kinetics of DNA methylation inheritance by the Dnmt1-including complexes during the cell cycle. *Cell division* 7:5.
- Hervouet E, Vallette FM and Cartron PF (2009) Dnmt3/transcription factor interactions as crucial players in targeted DNA methylation. *Epigenetics* 4:487-499.
- Hervouet E, Vallette FM and Cartron PF (2010b) Dnmt1/Transcription factor interactions: an alternative mechanism of DNA methylation inheritance. *Genes & cancer* 1:434-443.
- Hodges E, Molaro A, Dos Santos CO, Thekkat P, Song Q, Uren PJ, Park J, Butler J, Rafii S, McCombie WR, Smith AD and Hannon GJ (2011) Directional DNA methylation changes and complex intermediate states accompany lineage specificity in the adult hematopoietic compartment. *Molecular cell* 44:17-28.
- Hodges E, Smith AD, Kendall J, Xuan Z, Ravi K, Rooks M, Zhang MQ, Ye K, Bhattacharjee A, Brizuela L, McCombie WR, Wigler M, Hannon GJ and Hicks JB (2009) High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. *Genome research* 19:1593-1605.
- Holliday R and Pugh JE (1975) DNA modification mechanisms and gene activity during development. *Science (New York, NY)* 187:226-232.
- Holz-Schietinger C and Reich NO (2012) RNA modulation of the human DNA methyltransferase 3A. *Nucleic acids research*.
- Hopfner R, Mousli M, Garnier JM, Redon R, du Manoir S, Chatton B, Ghyselinck N, Oudet P and Bronner C (2001) Genomic structure and chromosomal mapping of the gene



- coding for ICBP90, a protein involved in the regulation of the topoisomerase IIalpha gene expression. *Gene* 266:15-23.
- Hopfner R, Mousli M, Jeltsch JM, Voulgaris A, Lutz Y, Marin C, Bellocq JP, Oudet P and Bronner C (2000) ICBP90, a novel human CCAAT binding protein, involved in the regulation of topoisomerase IIalpha expression. *Cancer research* 60:121-128.
- Hu L, Li Z, Wang P, Lin Y and Xu Y (2011) Crystal structure of PHD domain of UHRF1 and insights into recognition of unmodified histone H3 arginine residue 2. *Cell research* 21:1374-1378.
- Huang V, Place RF, Portnoy V, Wang J, Qi Z, Jia Z, Yu A, Shuman M, Yu J and Li LC (2012) Upregulation of Cyclin B1 by miRNA and its implications in cancer. *Nucleic acids research* 40:1695-1707.
- Huertas D, Sendra R and Munoz P (2009) Chromatin dynamics coupled to DNA repair. *Epigenetics* 4:31-42.
- Hung MS and Shen CK (2003) Eukaryotic methyl-CpG-binding domain proteins and chromatin modification. *Eukaryotic cell* 2:841-846.
- Hung T, Wang Y, Lin MF, Koegel AK, Kotake Y, Grant GD, Horlings HM, Shah N, Umbricht C, Wang P, Wang Y, Kong B, Langerod A, Borresen-Dale AL, Kim SK, van de Vijver M, Sukumar S, Whitfield ML, Kellis M, Xiong Y, Wong DJ and Chang HY (2011) Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nature genetics* 43:621-629.
- Hussain SP and Harris CC (2000) Molecular epidemiology and carcinogenesis: endogenous and exogenous carcinogens. *Mutation research* 462:311-322.
- Iliopoulos D, Lindahl-Allen M, Polytarchou C, Hirsch HA, Tschlis PN and Struhl K (2010) Loss of miR-200 inhibition of Suz12 leads to polycomb-mediated repression required for the formation and maintenance of cancer stem cells. *Molecular cell* 39:761-772.
- Illingworth R, Kerr A, Desousa D, Jorgensen H, Ellis P, Stalker J, Jackson D, Clee C, Plumb R, Rogers J, Humphray S, Cox T, Langford C and Bird A (2008) A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS biology* 6:e22.
- Illingworth RS and Bird AP (2009) CpG islands--'a rough guide'. *FEBS letters* 583:1713-1720.

- Imamura T, Ohgane J, Ito S, Ogawa T, Hattori N, Tanaka S and Shiota K (2001) CpG island of rat sphingosine kinase-1 gene: tissue-dependent DNA methylation status and multiple alternative first exons. *Genomics* 76:117-125.
- Inoue A and Zhang Y (2011) Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science (New York, NY)* 334:194.
- Iorio MV and Croce CM (2012) MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. *EMBO molecular medicine* 4:143-159.
- Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, Ji H, Potash JB, Sabunciyan S and Feinberg AP (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nature genetics* 41:178-186.
- Ishida M and Moore GE (2012) The role of imprinted genes in humans. *Molecular aspects of medicine*.
- Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C and Zhang Y (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science (New York, NY)* 333:1300-1303.
- Jair KW, Bachman KE, Suzuki H, Ting AH, Rhee I, Yen RW, Baylin SB and Schuebel KE (2006) De novo CpG island methylation in human cancer cells. *Cancer research* 66:682-692.
- Jeanblanc M, Mousli M, Hopfner R, Bathami K, Martinet N, Abbady AQ, Siffert JC, Mathieu E, Muller CD and Bronner C (2005) The retinoblastoma gene and its product are targeted by ICBP90: a key mechanism in the G1/S transition during the cell cycle. *Oncogene* 24:7337-7345.
- Jeffery L and Nakielny S (2004) Components of the DNA methylation system of chromatin control are RNA-binding proteins. *The Journal of biological chemistry* 279:49479-49487.
- Jenkins Y, Markovtsov V, Lang W, Sharma P, Pearsall D, Warner J, Franci C, Huang B, Huang J, Yam GC, Vistan JP, Pali E, Vialard J, Janicot M, Lorens JB, Payan DG and Hitoshi Y (2005) Critical role of the ubiquitin ligase activity of UHRF1, a nuclear RING finger protein, in tumor cell growth. *Molecular biology of the cell* 16:5621-5629.

- Jeong S, Liang G, Sharma S, Lin JC, Choi SH, Han H, Yoo CB, Egger G, Yang AS and Jones PA (2009) Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. *Molecular and cellular biology* 29:5366-5376.
- Ji AJ, Jiang Z, Livson Y, Davis JA, Chu JX and Weng N (2010) Challenges in urine bioanalytical assays: overcoming nonspecific binding. *Bioanalysis* 2:1573-1586.
- Jia D, Jurkowska RZ, Zhang X, Jeltsch A and Cheng X (2007) Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* 449:248-251.
- Jiang C and Pugh BF (2009) Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet* 10:161-172.
- Jones PA (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 13:484-492.
- Jones PA and Liang G (2009) Rethinking how DNA methylation patterns are maintained. *Nat Rev Genet* 10:805-811.
- Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J and Wolffe AP (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature genetics* 19:187-191.
- Juergens RA, Wrangle J, Vendetti FP, Murphy SC, Zhao M, Coleman B, Sebree R, Rodgers K, Hooker CM, Franco N, Lee B, Tsai S, Delgado IE, Rudek MA, Belinsky SA, Herman JG, Baylin SB, Brock MV and Rudin CM (2011) Combination epigenetic therapy has efficacy in patients with refractory advanced non-small cell lung cancer. *Cancer discovery* 1:598-607.
- Jurkowska RZ, Jurkowski TP and Jeltsch A (2011) Structure and function of mammalian DNA methyltransferases. *Chembiochem* 12:206-222.
- Kaati G, Bygren LO and Edvinsson S (2002) Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period. *Eur J Hum Genet* 10:682-688.
- Kamakaka RT and Biggins S (2005) Histone variants: deviants? *Genes & development* 19:295-310.
- Kaneda M, Hirasawa R, Chiba H, Okano M, Li E and Sasaki H (2010) Genetic evidence for Dnmt3a-dependent imprinting during oocyte growth obtained by conditional knockout with Zp3-Cre and complete exclusion of Dnmt3b by chimera formation. *Genes Cells* 15:169-179.

- Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E and Sasaki H (2004) Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* 429:900-903.
- Karagianni P, Amazit L, Qin J and Wong J (2008) ICBP90, a novel methyl K9 H3 binding protein linking protein ubiquitination with heterochromatin formation. *Molecular and cellular biology* 28:705-717.
- Kareta MS, Botello ZM, Ennis JJ, Chou C and Chedin F (2006) Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L. *The Journal of biological chemistry* 281:25893-25902.
- Kato Y, Kaneda M, Hata K, Kumaki K, Hisano M, Kohara Y, Okano M, Li E, Nozaki M and Sasaki H (2007) Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Human molecular genetics* 16:2272-2280.
- Kiefer CM, Hou C, Little JA and Dean A (2008) Epigenetics of beta-globin gene regulation. *Mutation research* 647:68-76.
- Kim DH, Saetrom P, Snove O, Jr. and Rossi JJ (2008) MicroRNA-directed transcriptional gene silencing in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* 105:16230-16235.
- Kim GD, Ni J, Kelesoglu N, Roberts RJ and Pradhan S (2002) Co-operation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases. *The EMBO journal* 21:4183-4195.
- Kim J, Guermah M, McGinty RK, Lee JS, Tang Z, Milne TA, Shilatifard A, Muir TW and Roeder RG (2009a) RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. *Cell* 137:459-471.
- Kim JK, Esteve PO, Jacobsen SE and Pradhan S (2009b) UHRF1 binds G9a and participates in p21 transcriptional regulation in mammalian cells. *Nucleic acids research* 37:493-505.
- Kim VN (2005) MicroRNA biogenesis: coordinated cropping and dicing. *Nature reviews* 6:376-385.
- Klimasauskas S, Kumar S, Roberts RJ and Cheng X (1994) HhaI methyltransferase flips its target base out of the DNA helix. *Cell* 76:357-369.

- Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS and Ahringer J (2009) Differential chromatin marking of introns and expressed exons by H3K36me3. *Nature genetics* 41:376-381.
- Kondo Y, Shen L, Cheng AS, Ahmed S, Bumber Y, Charo C, Yamochi T, Urano T, Furukawa K, Kwabi-Addo B, Gold DL, Sekido Y, Huang TH and Issa JP (2008) Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. *Nature genetics* 40:741-750.
- Koukoura O, Sifakis S and Spandidos DA (2012) DNA methylation in the human placenta and fetal growth (review). *Molecular medicine reports* 5:883-889.
- Kouzarides T (2007) Chromatin modifications and their function. *Cell* 128:693-705.
- Kriaucionis S and Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science (New York, NY)* 324:929-930.
- Kuramochi-Miyagawa S, Watanabe T, Gotoh K, Totoki Y, Toyoda A, Ikawa M, Asada N, Kojima K, Yamaguchi Y, Ijiri TW, Hata K, Li E, Matsuda Y, Kimura T, Okabe M, Sakaki Y, Sasaki H and Nakano T (2008) DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes & development* 22:908-917.
- Ladurner AG (2003) Inactivating chromosomes: a macro domain that minimizes transcription. *Molecular cell* 12:1-3.
- Lakowicz JR (1999) *Principles of fluorescence spectroscopy, Second edition*, Academic/Plenum Publishers, New York.
- Lakowicz JR and Maliwal BP (1985) Construction and performance of a variable-frequency phase-modulation fluorometer. *Biophysical chemistry* 21:61-78.
- Lallous N, Legrand P, McEwen AG, Ramon-Maiques S, Samama JP and Birck C (2011) The PHD finger of human UHRF1 reveals a new subgroup of unmethylated histone H3 tail readers. *PloS one* 6:e27599.
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al. (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860-921.
- Lane N, Dean W, Erhardt S, Hajkova P, Surani A, Walter J and Reik W (2003) Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* 35:88-93.

- Langevin SM, Houseman EA, Christensen BC, Wiencke JK, Nelson HH, Karagas MR, Marsit CJ and Kelsey KT (2011) The influence of aging, environmental exposures and local sequence features on the variation of DNA methylation in blood. *Epigenetics* 6:908-919.
- Laurent L, Wong E, Li G, Huynh T, Tsigos A, Ong CT, Low HM, Kin Sung KW, Rigoutsos I, Loring J and Wei CL (2010) Dynamic changes in the human methylome during differentiation. *Genome research* 20:320-331.
- Lee JS, Shukla A, Schneider J, Swanson SK, Washburn MP, Florens L, Bhaumik SR and Shilatifard A (2007) Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. *Cell* 131:1084-1096.
- Lee JS, Smith E and Shilatifard A (2010) The language of histone crosstalk. *Cell* 142:682-685.
- Lee KK and Workman JL (2007) Histone acetyltransferase complexes: one size doesn't fit all. *Nature reviews* 8:284-295.
- Lee TT, Agarwalla S and Stroud RM (2005) A unique RNA Fold in the RuvA-RNA-cofactor ternary complex contributes to substrate selectivity and enzymatic function. *Cell* 120:599-611.
- Lee YS and Dutta A (2009) MicroRNAs in cancer. *Annual review of pathology* 4:199-227.
- Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T and Peters AH (2003) Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* 13:1192-1200.
- Leonhardt H, Page AW, Weier HU and Bestor TH (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71:865-873.
- Lewis A and Reik W (2006) How imprinting centres work. *Cytogenetic and genome research* 113:81-89.
- Li B, Carey M and Workman JL (2007) The role of chromatin during transcription. *Cell* 128:707-719.
- Li E, Bestor TH and Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69:915-926.
- Li G and Reinberg D (2011) Chromatin higher-order structures and gene regulation. *Current opinion in genetics & development* 21:175-186.



- Li Y, Mori T, Hata H, Homma Y and Kochi H (2004) NIRF induces G1 arrest and associates with Cdk2. *Biochemical and biophysical research communications* 319:464-468.
- Liang G, Chan MF, Tomigahara Y, Tsai YC, Gonzales FA, Li E, Laird PW and Jones PA (2002) Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Molecular and cellular biology* 22:480-491.
- Liang G, Lin JC, Wei V, Yoo C, Cheng JC, Nguyen CT, Weisenberger DJ, Egger G, Takai D, Gonzales FA and Jones PA (2004) Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. *Proceedings of the National Academy of Sciences of the United States of America* 101:7357-7362.
- Lim DH and Maher ER (2010) Genomic imprinting syndromes and cancer. *Advances in genetics* 70:145-175.
- Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B and Ecker JR (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462:315-322.
- Liu S, Abbondanzieri EA, Rausch JW, Le Grice SF and Zhuang X (2008) Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. *Science (New York, NY)* 322:1092-1097.
- Lorenzato M, Caudroy S, Bronner C, Evrard G, Simon M, Durlach A, Birembaut P and Clavel C (2005) Cell cycle and/or proliferation markers: what is the best method to discriminate cervical high-grade lesions? *Human pathology* 36:1101-1107.
- Lucifero D, La Salle S, Bourc'his D, Martel J, Bestor TH and Trasler JM (2007) Coordinate regulation of DNA methyltransferase expression during oogenesis. *BMC developmental biology* 7:36.
- Lucifero D, Mann MR, Bartolomei MS and Trasler JM (2004) Gene-specific timing and epigenetic memory in oocyte imprinting. *Human molecular genetics* 13:839-849.
- Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM and Misteli T (2010) Regulation of alternative splicing by histone modifications. *Science (New York, NY)* 327:996-1000.
- Luger K and Hansen JC (2005) Nucleosome and chromatin fiber dynamics. *Current opinion in structural biology* 15:188-196.
- Luger K, Mader AW, Richmond RK, Sargent DF and Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389:251-260.

- Lumey LH, Stein AD, Kahn HS and Romijn JA (2009) Lipid profiles in middle-aged men and women after famine exposure during gestation: the Dutch Hunger Winter Families Study. *The American journal of clinical nutrition* 89:1737-1743.
- Ma H, Chen H, Guo X, Wang Z, Sowa ME, Zheng L, Hu S, Zeng P, Guo R, Diao J, Lan F, Harper JW, Shi YG, Xu Y and Shi Y (2012) M phase phosphorylation of the epigenetic regulator UHRF1 regulates its physical association with the deubiquitylase USP7 and stability. *Proceedings of the National Academy of Sciences of the United States of America* 109:4828-4833.
- Maegawa S, Hinkal G, Kim HS, Shen L, Zhang L, Zhang J, Zhang N, Liang S, Donehower LA and Issa JP (2010) Widespread and tissue specific age-related DNA methylation changes in mice. *Genome research* 20:332-340.
- Maiti A and Drohat AC (2011) Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *The Journal of biological chemistry* 286:35334-35338.
- Malumbres M (2012) miRNAs and cancer: An epigenetics view. *Molecular aspects of medicine*.
- Margot JB, Ehrenhofer-Murray AE and Leonhardt H (2003) Interactions within the mammalian DNA methyltransferase family. *BMC molecular biology* 4:7.
- Matei DE and Nephew KP (2010) Epigenetic therapies for chemoresensitization of epithelial ovarian cancer. *Gynecologic oncology* 116:195-201.
- Mattick JS and Makunin IV (2005) Small regulatory RNAs in mammals. *Human molecular genetics* 14 Spec No 1:R121-132.
- Mayer W, Niveleau A, Walter J, Fundele R and Haaf T (2000) Demethylation of the zygotic paternal genome. *Nature* 403:501-502.
- McKittrick E, Gafken PR, Ahmad K and Henikoff S (2004) Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proceedings of the National Academy of Sciences of the United States of America* 101:1525-1530.
- Medina-Franco JL and Caulfield T (2011) Advances in the computational development of DNA methyltransferase inhibitors. *Drug discovery today* 16:418-425.
- Meilinger D, Fellingner K, Bultmann S, Rothbauer U, Bonapace IM, Klinkert WE, Spada F and Leonhardt H (2009) Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells. *EMBO reports* 10:1259-1264.

- Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R and Lander ES (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454:766-770.
- Mely Y, Jullian N, Morellet N, De Rocquigny H, Dong CZ, Piemont E, Roques BP and Gerard D (1994) Spatial proximity of the HIV-1 nucleocapsid protein zinc fingers investigated by time-resolved fluorescence and fluorescence resonance energy transfer. *Biochemistry* 33:12085-12091.
- Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES and Bernstein BE (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448:553-560.
- Mistry H, Gibson L, Yun JW, Sarras H, Tamblyn L and McPherson JP (2008) Interplay between Np95 and Eme1 in the DNA damage response. *Biochemical and biophysical research communications* 375:321-325.
- Mohammad F, Mondal T and Kanduri C (2009) Epigenetics of imprinted long noncoding RNAs. *Epigenetics* 4:277-286.
- Morgan HD, Santos F, Green K, Dean W and Reik W (2005) Epigenetic reprogramming in mammals. *Human molecular genetics* 14 Spec No 1:R47-58.
- Mori T, Ikeda DD, Fukushima T, Takenoshita S and Kochi H (2011) NIRF constitutes a nodal point in the cell cycle network and is a candidate tumor suppressor. *Cell cycle (Georgetown, Tex)* 10:3284-3299.
- Mori T, Ikeda DD, Yamaguchi Y and Unoki M (2012) NIRF/UHRF2 occupies a central position in the cell cycle network and allows coupling with the epigenetic landscape. *FEBS letters* 586:1570-1583.
- Mousli M, Hopfner R, Abbady AQ, Monte D, Jeanblanc M, Oudet P, Louis B and Bronner C (2003) ICBP90 belongs to a new family of proteins with an expression that is deregulated in cancer cells. *British journal of cancer* 89:120-127.
- Muto M, Kanari Y, Kubo E, Takabe T, Kurihara T, Fujimori A and Tatsumi K (2002) Targeted disruption of Np95 gene renders murine embryonic stem cells hypersensitive to DNA damaging agents and DNA replication blocks. *The Journal of biological chemistry* 277:34549-34555.

- Nady N, Lemak A, Walker JR, Avvakumov GV, Kareta MS, Achour M, Xue S, Duan S, Allali-Hassani A, Zuo X, Wang YX, Bronner C, Chedin F, Arrowsmith CH and Dhe-Paganon S (2011) Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein. *The Journal of biological chemistry* 286:24300-24311.
- Nakanishi S, Lee JS, Gardner KE, Gardner JM, Takahashi YH, Chandrasekharan MB, Sun ZW, Osley MA, Strahl BD, Jaspersen SL and Shilatifard A (2009) Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1. *The Journal of cell biology* 186:371-377.
- Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN and Bird A (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393:386-389.
- Ng SS, Yue WW, Oppermann U and Klose RJ (2009) Dynamic protein methylation in chromatin biology. *Cell Mol Life Sci* 66:407-422.
- Noonan EJ, Place RF, Pookot D, Basak S, Whitson JM, Hirata H, Giardina C and Dahiya R (2009) miR-449a targets HDAC-1 and induces growth arrest in prostate cancer. *Oncogene* 28:1714-1724.
- Nordlund TM, Xu D and Evans KO (1993) Excitation energy transfer in DNA: duplex melting and transfer from normal bases to 2-aminopurine. *Biochemistry* 32:12090-12095.
- Oba-Shinjo SM, Bengtson MH, Winnischofer SM, Colin C, Vedoy CG, de Mendonca Z, Marie SK and Sogayar MC (2005) Identification of novel differentially expressed genes in human astrocytomas by cDNA representational difference analysis. *Brain research* 140:25-33.
- Okano M, Bell DW, Haber DA and Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99:247-257.
- Okano M, Xie S and Li E (1998) Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature genetics* 19:219-220.
- Oki M, Aihara H and Ito T (2007) Role of histone phosphorylation in chromatin dynamics and its implications in diseases. *Sub-cellular biochemistry* 41:319-336.
- Olsen EA, Kim YH, Kuzel TM, Pacheco TR, Foss FM, Parker S, Frankel SR, Chen C, Ricker JL, Arduino JM and Duvic M (2007) Phase IIb multicenter trial of vorinostat in

- patients with persistent, progressive, or treatment refractory cutaneous T-cell lymphoma. *J Clin Oncol* 25:3109-3115.
- Ooi SK and Bestor TH (2008) The colorful history of active DNA demethylation. *Cell* 133:1145-1148.
- Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, Cheng X and Bestor TH (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* 448:714-717.
- Otani J, Nankumo T, Arita K, Inamoto S, Ariyoshi M and Shirakawa M (2009) Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO reports* 10:1235-1241.
- Ozdag H, Teschendorff AE, Ahmed AA, Hyland SJ, Blenkiron C, Bobrow L, Veerakumarasivam A, Burt G, Subkhankulova T, Arends MJ, Collins VP, Bowtell D, Kouzarides T, Brenton JD and Caldas C (2006) Differential expression of selected histone modifier genes in human solid cancers. *BMC genomics* 7:90.
- Papait R, Pistore C, Grazini U, Babbio F, Cogliati S, Pecoraro D, Brino L, Morand AL, Dechampsme AM, Spada F, Leonhardt H, McBlane F, Oudet P and Bonapace IM (2008) The PHD domain of Np95 (mUHRF1) is involved in large-scale reorganization of pericentromeric heterochromatin. *Molecular biology of the cell* 19:3554-3563.
- Parker JB, Bianchet MA, Krosky DJ, Friedman JI, Amzel LM and Stivers JT (2007) Enzymatic capture of an extrahelical thymine in the search for uracil in DNA. *Nature* 449:433-437.
- Parry L and Clarke AR (2011) The Roles of the Methyl-CpG Binding Proteins in Cancer. *Genes & cancer* 2:618-630.
- Pembrey ME, Bygren LO, Kaati G, Edvinsson S, Northstone K, Sjöström M and Golding J (2006) Sex-specific, male-line transgenerational responses in humans. *Eur J Hum Genet* 14:159-166.
- Pfaffeneder T, Hackner B, Truss M, Munzel M, Müller M, Deiml CA, Hagemeyer C and Carell T (2011) The discovery of 5-formylcytosine in embryonic stem cell DNA. *Angewandte Chemie (International ed)* 50:7008-7012.
- Pfeifer GP, Tang M and Denissenko MF (2000) Mutation hotspots and DNA methylation. *Current topics in microbiology and immunology* 249:1-19.
- Pichler G, Wolf P, Schmidt CS, Meilinger D, Schneider K, Frauer C, Fellingner K, Rottach A and Leonhardt H (2011) Cooperative DNA and histone binding by Uhrf2 links the two

- major repressive epigenetic pathways. *Journal of cellular biochemistry* 112:2585-2593.
- Piekarz RL, Frye R, Turner M, Wright JJ, Allen SL, Kirschbaum MH, Zain J, Prince HM, Leonard JP, Geskin LJ, Reeder C, Joske D, Figg WD, Gardner ER, Steinberg SM, Jaffe ES, Stetler-Stevenson M, Lade S, Fojo AT and Bates SE (2009) Phase II multi-institutional trial of the histone deacetylase inhibitor romidepsin as monotherapy for patients with cutaneous T-cell lymphoma. *J Clin Oncol* 27:5410-5417.
- Place RF, Li LC, Pookot D, Noonan EJ and Dahiya R (2008) MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proceedings of the National Academy of Sciences of the United States of America* 105:1608-1613.
- Pontier DB and Gribnau J (2011) Xist regulation and function explored. *Human genetics* 130:223-236.
- Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, Jacobsen SE and Reik W (2010) Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* 463:1101-1105.
- Portela A and Esteller M (2010) Epigenetic modifications and human disease. *Nature biotechnology* 28:1057-1068.
- Prokhortchouk A, Hendrich B, Jorgensen H, Ruzov A, Wilm M, Georgiev G, Bird A and Prokhortchouk E (2001) The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes & development* 15:1613-1618.
- Qian C, Li S, Jakoncic J, Zeng L, Walsh MJ and Zhou MM (2008) Structure and hemimethylated CpG binding of the SRA domain from human UHRF1. *The Journal of biological chemistry* 283:34490-34494.
- Qin W, Leonhardt H and Spada F (2011) Usp7 and Uhrf1 control ubiquitination and stability of the maintenance DNA methyltransferase Dnmt1. *Journal of cellular biochemistry* 112:439-444.
- Rajakumara E, Wang Z, Ma H, Hu L, Chen H, Lin Y, Guo R, Wu F, Li H, Lan F, Shi YG, Xu Y, Patel DJ and Shi Y (2011) PHD finger recognition of unmodified histone H3R2 links UHRF1 to regulation of euchromatic gene expression. *Molecular cell* 43:275-284.
- Ramsahoye BH, Biniszkiwicz D, Lyko F, Clark V, Bird AP and Jaenisch R (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA



- methyltransferase 3a. *Proceedings of the National Academy of Sciences of the United States of America* 97:5237-5242.
- Rando OJ and Chang HY (2009) Genome-wide views of chromatin structure. *Annual review of biochemistry* 78:245-271.
- Reik W, Dean W and Walter J (2001) Epigenetic reprogramming in mammalian development. *Science (New York, NY)* 293:1089-1093.
- Reik W and Walter J (2001) Genomic imprinting: parental influence on the genome. *Nat Rev Genet* 2:21-32.
- Riggs AD (1975) X inactivation, differentiation, and DNA methylation. *Cytogenetics and cell genetics* 14:9-25.
- Robertson KD, Ait-Si-Ali S, Yokochi T, Wade PA, Jones PL and Wolffe AP (2000a) DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nature genetics* 25:338-342.
- Robertson KD, Keyomarsi K, Gonzales FA, Velicescu M and Jones PA (2000b) Differential mRNA expression of the human DNA methyltransferases (DNMTs) 1, 3a and 3b during the G(0)/G(1) to S phase transition in normal and tumor cells. *Nucleic acids research* 28:2108-2113.
- Robinson PJ, Fairall L, Huynh VA and Rhodes D (2006) EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. *Proceedings of the National Academy of Sciences of the United States of America* 103:6506-6511.
- Rodriguez-Paredes M and Esteller M (2011) Cancer epigenetics reaches mainstream oncology. *Nature medicine* 17:330-339.
- Rosenfeld CS (2010) Animal models to study environmental epigenetics. *Biology of reproduction* 82:473-488.
- Rothbart SB, Krajewski K, Nady N, Tempel W, Xue S, Badeaux AI, Barsyte-Lovejoy D, Martinez JY, Bedford MT, Fuchs SM, Arrowsmith CH and Strahl BD (2012) Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nature structural & molecular biology*.
- Rottach A, Frauer C, Pichler G, Bonapace IM, Spada F and Leonhardt H (2010) The multi-domain protein Np95 connects DNA methylation and histone modification. *Nucleic acids research* 38:1796-1804.

- Rottach A, Leonhardt H and Spada F (2009) DNA methylation-mediated epigenetic control. *Journal of cellular biochemistry* 108:43-51.
- Sabatino L, Fucci A, Pancione M, Carafa V, Nebbioso A, Pistore C, Babbio F, Votino C, Laudanna C, Ceccarelli M, Altucci L, Bonapace IM and Colantuoni V (2012) UHRF1 coordinates peroxisome proliferator activated receptor gamma (PPARG) epigenetic silencing and mediates colorectal cancer progression. *Oncogene*.
- Sandoval J and Esteller M (2012) Cancer epigenomics: beyond genomics. *Current opinion in genetics & development* 22:50-55.
- Santenard A and Torres-Padilla ME (2009) Epigenetic reprogramming in mammalian reproduction: contribution from histone variants. *Epigenetics* 4:80-84.
- Sasai N and Defossez PA (2009) Many paths to one goal? The proteins that recognize methylated DNA in eukaryotes. *The International journal of developmental biology* 53:323-334.
- Saxonov S, Berg P and Brutlag DL (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proceedings of the National Academy of Sciences of the United States of America* 103:1412-1417.
- Schaaf GJ, Ruijter JM, van Ruissen F, Zwijnenburg DA, Waaijer R, Valentijn LJ, Benit-Deekman J, van Kampen AH, Baas F and Kool M (2005) Full transcriptome analysis of rhabdomyosarcoma, normal, and fetal skeletal muscle: statistical comparison of multiple SAGE libraries. *Faseb J* 19:404-406.
- Schalch T, Duda S, Sargent DF and Richmond TJ (2005) X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* 436:138-141.
- Schermelleh L, Haemmer A, Spada F, Rosing N, Meilinger D, Rothbauer U, Cardoso MC and Leonhardt H (2007) Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. *Nucleic acids research* 35:4301-4312.
- Schilling E and Rehli M (2007) Global, comparative analysis of tissue-specific promoter CpG methylation. *Genomics* 90:314-323.
- Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J, Eden E, Yakhini Z, Ben-Shushan E, Reubinoff BE, Bergman Y, Simon I and Cedar H (2007) Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. *Nature genetics* 39:232-236.

- Schlick T, Hayes J and Grigoryev S (2012) Toward convergence of experimental studies and theoretical modeling of the chromatin fiber. *The Journal of biological chemistry* 287:5183-5191.
- Schmitz KM, Mayer C, Postepska A and Grummt I (2010) Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes & development* 24:2264-2269.
- Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G and Zhao K (2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell* 132:887-898.
- Schultz DC, Ayyanathan K, Negorev D, Maul GG and Rauscher FJ, 3rd (2002) SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes & development* 16:919-932.
- Shahbazian MD and Grunstein M (2007) Functions of site-specific histone acetylation and deacetylation. *Annual review of biochemistry* 76:75-100.
- Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K, Tajima S, Mitsuya K, Okano M and Koseki H (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* 450:908-912.
- Sharma S, Kelly TK and Jones PA (2010) Epigenetics in cancer. *Carcinogenesis* 31:27-36.
- Shen L, Kondo Y, Guo Y, Zhang J, Zhang L, Ahmed S, Shu J, Chen X, Waterland RA and Issa JP (2007) Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. *PLoS genetics* 3:2023-2036.
- Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA and Shi Y (2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119:941-953.
- Shih TY and Fasman GD (1972) Circular dichroism studies of histone-deoxyribonucleic acid complexes. A comparison of complexes with histone I (f-1), histone IV (f2al), and their mixtures. *Biochemistry* 11:398-404.
- Shiio Y and Eisenman RN (2003) Histone sumoylation is associated with transcriptional repression. *Proceedings of the National Academy of Sciences of the United States of America* 100:13225-13230.

- Shiota K, Kogo Y, Ohgane J, Imamura T, Urano A, Nishino K, Tanaka S and Hattori N (2002) Epigenetic marks by DNA methylation specific to stem, germ and somatic cells in mice. *Genes Cells* 7:961-969.
- Smallwood A, Esteve PO, Pradhan S and Carey M (2007) Functional cooperation between HP1 and DNMT1 mediates gene silencing. *Genes & development* 21:1169-1178.
- Smallwood SA and Kelsey G (2012) De novo DNA methylation: a germ cell perspective. *Trends Genet* 28:33-42.
- Smallwood SA, Tomizawa S, Krueger F, Ruf N, Carli N, Segonds-Pichon A, Sato S, Hata K, Andrews SR and Kelsey G (2011) Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nature genetics* 43:811-814.
- Smith CL and Peterson CL (2005) ATP-dependent chromatin remodeling. *Current topics in developmental biology* 65:115-148.
- Song F, Smith JF, Kimura MT, Morrow AD, Matsuyama T, Nagase H and Held WA (2005) Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proceedings of the National Academy of Sciences of the United States of America* 102:3336-3341.
- Stein IH, Schuller V, Bohm P, Tinnefeld P and Liedl T (2011) Single-molecule FRET ruler based on rigid DNA origami blocks. *Chemphyschem* 12:689-695.
- Strahl BD and Allis CD (2000) The language of covalent histone modifications. *Nature* 403:41-45.
- Straussman R, Nejman D, Roberts D, Steinfeld I, Blum B, Benvenisty N, Simon I, Yakhini Z and Cedar H (2009) Developmental programming of CpG island methylation profiles in the human genome. *Nature structural & molecular biology* 16:564-571.
- Stresemann C and Lyko F (2008) Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine. *International journal of cancer* 123:8-13.
- Su Z, Xia J and Zhao Z (2011) Functional complementation between transcriptional methylation regulation and post-transcriptional microRNA regulation in the human genome. *BMC genomics* 12 Suppl 5:S15.
- Suetake I, Shinozaki F, Miyagawa J, Takeshima H and Tajima S (2004) DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. *The Journal of biological chemistry* 279:27816-27823.

- Svotelis A, Gevry N and Gaudreau L (2009) Regulation of gene expression and cellular proliferation by histone H2A.Z. *Biochemistry and cell biology = Biochimie et biologie cellulaire* 87:179-188.
- Tachibana M, Matsumura Y, Fukuda M, Kimura H and Shinkai Y (2008) G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *The EMBO journal* 27:2681-2690.
- Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L and Rao A (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science (New York, NY)* 324:930-935.
- Talbert PB and Henikoff S (2010) Histone variants--ancient wrap artists of the epigenome. *Nature reviews* 11:264-275.
- Tien AL, Senbanerjee S, Kulkarni A, Mudbhary R, Goudreau B, Ganesan S, Sadler KC and Ukomadu C (2011) UHRF1 depletion causes a G2/M arrest, activation of DNA damage response and apoptosis. *The Biochemical journal* 435:175-185.
- Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD, Slagboom PE and Heijmans BT (2009) DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Human molecular genetics* 18:4046-4053.
- Tollervey JR and Lunyak VV (2012) Epigenetics: Judge, jury and executioner of stem cell fate. *Epigenetics* 7:823-840.
- Tremblay KD, Saam JR, Ingram RS, Tilghman SM and Bartolomei MS (1995) A paternal-specific methylation imprint marks the alleles of the mouse H19 gene. *Nature genetics* 9:407-413.
- Tremethick DJ (2007) Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell* 128:651-654.
- Trojer P and Reinberg D (2007) Facultative heterochromatin: is there a distinctive molecular signature? *Molecular cell* 28:1-13.
- Trotzler MA, Bronner C, Bathami K, Mathieu E, Abbady AQ, Jeanblanc M, Muller CD, Rochette-Egly C and Mousli M (2004) Phosphorylation of ICBP90 by protein kinase A enhances topoisomerase IIalpha expression. *Biochemical and biophysical research communications* 319:590-595.
- Tsai HC, Li H, Van Neste L, Cai Y, Robert C, Rassool FV, Shin JJ, Harbom KM, Beaty R, Pappou E, Harris J, Yen RW, Ahuja N, Brock MV, Stearns V, Feller-Kopman D,

- Yarmus LB, Lin YC, Welm AL, Issa JP, Minn I, Matsui W, Jang YY, Sharkis SJ, Baylin SB and Zahnow CA (2012) Transient low doses of DNA-demethylating agents exert durable antitumor effects on hematological and epithelial tumor cells. *Cancer cell* 21:430-446.
- Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P and Zhang Y (2006) Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 439:811-816.
- Turner JM, Mahadevaiah SK, Fernandez-Capetillo O, Nussenzweig A, Xu X, Deng CX and Burgoyne PS (2005) Silencing of unsynapsed meiotic chromosomes in the mouse. *Nature genetics* 37:41-47.
- Uemura T, Kubo E, Kanari Y, Ikemura T, Tatsumi K and Muto M (2000) Temporal and spatial localization of novel nuclear protein NP95 in mitotic and meiotic cells. *Cell structure and function* 25:149-159.
- Un F, Qi C, Prosser M, Wang N, Zhou B, Bronner C and Yen Y (2006) Modulating ICBP90 to suppress human ribonucleotide reductase M2 induction restores sensitivity to hydroxyurea cytotoxicity. *Anticancer research* 26:2761-2767.
- Unoki M, Brunet J and Mousli M (2009a) Drug discovery targeting epigenetic codes: the great potential of UHRF1, which links DNA methylation and histone modifications, as a drug target in cancers and toxoplasmosis. *Biochemical pharmacology* 78:1279-1288.
- Unoki M, Daigo Y, Koinuma J, Tsuchiya E, Hamamoto R and Nakamura Y (2010) UHRF1 is a novel diagnostic marker of lung cancer. *British journal of cancer* 103:217-222.
- Unoki M, Kelly JD, Neal DE, Ponder BA, Nakamura Y and Hamamoto R (2009b) UHRF1 is a novel molecular marker for diagnosis and the prognosis of bladder cancer. *British journal of cancer* 101:98-105.
- Unoki M, Nishidate T and Nakamura Y (2004) ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. *Oncogene* 23:7601-7610.
- Urduingio RG, Sanchez-Mut JV and Esteller M (2009) Epigenetic mechanisms in neurological diseases: genes, syndromes, and therapies. *Lancet neurology* 8:1056-1072.
- Verdel A, Vavasseur A, Le Gorrec M and Touat-Todeschini L (2009) Common themes in siRNA-mediated epigenetic silencing pathways. *The International journal of developmental biology* 53:245-257.



- Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG and Mann M (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142:967-980.
- Vilkaitis G, Suetake I, Klimasauskas S and Tajima S (2005) Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase. *The Journal of biological chemistry* 280:64-72.
- Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden JM, Bollen M, Esteller M, Di Croce L, de Launoit Y and Fuks F (2006) The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* 439:871-874.
- Vuilleumier C, Maechling-Strasser C, Gerard D and Mely Y (1997) Evidence and prevention of HIV-1 nucleocapsid protein adsorption onto fluorescence quartz cells. *Anal Biochem* 244:183-185.
- Walker JR, Avvakumov GV, Xue S, Dong A, Li Y, Bountra C, Weigelt J, Arrowsmith CH, Edwards AM, Bochkarev A and Dhe-Paganon S (2008) Cryptic tandem tudor domains in UHRF1 interact with H3K9Me and are important for pericentric heterochromatin replication, in (*PDB: 3db3*).
- Wang F, Yang YZ, Shi CZ, Zhang P, Moyer MP, Zhang HZ, Zou Y and Qin HL (2012) UHRF1 promotes cell growth and metastasis through repression of p16(ink4a) in colorectal cancer. *Annals of surgical oncology* 19:2753-2762.
- Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS and Zhang Y (2004) Role of histone H2A ubiquitination in Polycomb silencing. *Nature* 431:873-878.
- Wang J, Hevi S, Kurash JK, Lei H, Gay F, Bajko J, Su H, Sun W, Chang H, Xu G, Gaudet F, Li E and Chen T (2009) The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nature genetics* 41:125-129.
- Wang KC and Chang HY (2011) Molecular mechanisms of long noncoding RNAs. *Molecular cell* 43:904-914.
- Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh TY, Peng W, Zhang MQ and Zhao K (2008) Combinatorial patterns of histone acetylations and methylations in the human genome. *Nature genetics* 40:897-903.
- Waterland RA, Kellermayer R, Laritsky E, Rayco-Solon P, Harris RA, Travisano M, Zhang W, Torskaya MS, Zhang J, Shen L, Manary MJ and Prentice AM (2010) Season of

- conception in rural gambia affects DNA methylation at putative human metastable epialleles. *PLoS genetics* 6:e1001252.
- Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M and Schubeler D (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature genetics* 39:457-466.
- Weinberg MS, Villeneuve LM, Ehsani A, Amarzguioui M, Aagaard L, Chen ZX, Riggs AD, Rossi JJ and Morris KV (2006) The antisense strand of small interfering RNAs directs histone methylation and transcriptional gene silencing in human cells. *RNA (New York, NY)* 12:256-262.
- Wellner U, Schubert J, Burk UC, Schmalhofer O, Zhu F, Sonntag A, Waldvogel B, Vannier C, Darling D, zur Hausen A, Brunton VG, Morton J, Sansom O, Schuler J, Stemmler MP, Herzberger C, Hopt U, Keck T, Brabletz S and Brabletz T (2009) The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nature cell biology* 11:1487-1495.
- Whitehead J, Pandey GK and Kanduri C (2009) Regulation of the mammalian epigenome by long noncoding RNAs. *Biochimica et biophysica acta* 1790:936-947.
- Wong JJ, Hawkins NJ and Ward RL (2007) Colorectal cancer: a model for epigenetic tumorigenesis. *Gut* 56:140-148.
- Wood AJ and Oakey RJ (2006) Genomic imprinting in mammals: emerging themes and established theories. *PLoS genetics* 2:e147.
- Woodcock CL and Ghosh RP (2010) Chromatin higher-order structure and dynamics. *Cold Spring Harbor perspectives in biology* 2:a000596.
- Wu H, Caffo B, Jaffee HA, Irizarry RA and Feinberg AP (2010a) Redefining CpG islands using hidden Markov models. *Biostatistics (Oxford, England)* 11:499-514.
- Wu H, Coskun V, Tao J, Xie W, Ge W, Yoshikawa K, Li E, Zhang Y and Sun YE (2010b) Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science (New York, NY)* 329:444-448.
- Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K, Sun YE and Zhang Y (2011) Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes & development* 25:679-684.
- Wu H and Zhang Y (2011) Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes & development* 25:2436-2452.

- Xie S, Jakoncic J and Qian C (2012) UHRF1 double tudor domain and the adjacent PHD finger act together to recognize K9me3-containing histone H3 tail. *Journal of molecular biology* 415:318-328.
- Xu DG and Nordlund TM (2000) Sequence dependence of energy transfer in DNA oligonucleotides. *Biophysical journal* 78:1042-1058.
- Yang CG, Yi C, Duguid EM, Sullivan CT, Jian X, Rice PA and He C (2008) Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA. *Nature* 452:961-965.
- Yang GL, Zhang LH, Bo JJ, Chen HG, Cao M, Liu DM and Huang YR (2012) UHRF1 is associated with tumor recurrence in non-muscle-invasive bladder cancer. *Medical oncology (Northwood, London, England)* 29:842-847.
- Yang JS and Lai EC (2011) Alternative miRNA biogenesis pathways and the interpretation of core miRNA pathway mutants. *Molecular cell* 43:892-903.
- Yang XJ and Seto E (2007) HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene* 26:5310-5318.
- You JS and Jones PA (2012) Cancer genetics and epigenetics: two sides of the same coin? *Cancer cell* 22:9-20.
- Younger ST and Corey DR (2011) Transcriptional gene silencing in mammalian cells by miRNA mimics that target gene promoters. *Nucleic acids research* 39:5682-5691.
- Yuan GC, Liu YJ, Dion MF, Slack MD, Wu LF, Altschuler SJ and Rando OJ (2005) Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science (New York, NY)* 309:626-630.
- Zardo G, Ciolfi A, Vian L, Starnes LM, Billi M, Racanicchi S, Maresca C, Fazi F, Travaglini L, Noguera N, Mancini M, Nanni M, Cimino G, Lo-Coco F, Grignani F and Nervi C (2012) Polycombs and microRNA-223 regulate human granulopoiesis by transcriptional control of target gene expression. *Blood* 119:4034-4046.
- Zhang J, Gao Q, Li P, Liu X, Jia Y, Wu W, Li J, Dong S, Koseki H and Wong J (2011) S phase-dependent interaction with DNMT1 dictates the role of UHRF1 but not UHRF2 in DNA methylation maintenance. *Cell research* 21:1723-1739.
- Zhang R, Zhang L and Yu W (2012) Genome-wide expression of non-coding RNA and global chromatin modification. *Acta biochimica et biophysica Sinica* 44:40-47.
- Zhang Y, Jurkowska R, Soeroes S, Rajavelu A, Dhayalan A, Bock I, Rathert P, Brandt O, Reinhardt R, Fischle W and Jeltsch A (2010) Chromatin methylation activity of

- Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic acids research* 38:4246-4253.
- Zhang Y, Rohde C, Tierling S, Jurkowski TP, Bock C, Santacruz D, Ragozin S, Reinhardt R, Groth M, Walter J and Jeltsch A (2009) DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. *PLoS genetics* 5:e1000438.
- Zhao Q, Rank G, Tan YT, Li H, Moritz RL, Simpson RJ, Cerruti L, Curtis DJ, Patel DJ, Allis CD, Cunningham JM and Jane SM (2009) PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. *Nature structural & molecular biology* 16:304-311.
- Zhou Y, Santoro R and Grummt I (2002) The chromatin remodeling complex NoRC targets HDAC1 to the ribosomal gene promoter and represses RNA polymerase I transcription. *The EMBO journal* 21:4632-4640.
- Zilberman D, Coleman-Derr D, Ballinger T and Henikoff S (2008) Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* 456:125-129.
- Zlatanova J and Thakar A (2008) H2A.Z: view from the top. *Structure* 16:166-179.

# Epigénétique et méthylation de l'ADN : Etude des mécanismes d'interaction du domaine SRA de UHRF1 avec l'ADN hémi-méthylé

## Résumé

La protéine UHRF1 est impliquée dans le maintien et la transmission des modifications épigénétiques. Lors du processus de réplication, elle recrute la méthyltransférase de l'ADN Dnmt1 au niveau des sites CpG hémi-méthylés via son domaine SRA (SET and RING Associated), favorisant la duplication des profils de méthylation. La structure tridimensionnelle du complexe SRA/ADN révèle que la protéine induit un basculement de la méthylcytosine qui permet un ancrage spécifique de la protéine sur les sites hémi-méthylés, facilitant le recrutement de la Dnmt1 au niveau de ces positions stratégiques. Dans ce contexte, notre projet vise à comprendre les mécanismes d'interaction du domaine SRA de UHRF1 avec l'ADN hémi-méthylé. Des oligonucléotides doubles brins ont été marqués à la 2-aminopurine, un analogue nucléosidique fluorescent sensible à l'environnement, à différentes positions au voisinage d'un unique site de reconnaissance CpG hémi-méthylé. Les mesures de spectroscopie de fluorescence à l'état stationnaire et résolues en temps de ces duplexes liés au domaine SRA nous ont permis de caractériser de manière site-spécifique les changements conformationnels induits par la liaison du domaine SRA. En accord avec la structure tridimensionnelle du complexe SRA/ADN, nos données suggèrent que le domaine SRA est capable de basculer la méthylcytosine tout en préservant la structure des autres bases dans le duplexe. Le domaine SRA semble se lier selon le même mécanisme aux duplexes hémi-méthylés, bi-méthylés et non-méthylés. La protéine UHRF1 jouerait ainsi un rôle de "lecteur" capable de scanner la séquence d'ADN à la recherche de sites hémi-méthylés.

Mots-clés : épigénétique ; méthylation de l'ADN ; protéine UHRF1 ; domaine SRA ; 2-aminopurine ; spectroscopie de fluorescence

## Résumé en anglais

The UHRF1 protein plays a key role in the maintenance and transmission of epigenetic modifications. During the replication process, it recruits the DNA methyltransferase Dnmt1 to hemi-methylated CpG sites via its SRA (SET and RING Associated) domain, promoting the duplication of the methylation profiles. The tridimensional structure of the SRA/DNA complex revealed that the protein induces a base-flipping of the methylcytosine that enables a specific anchoring of the protein to hemi-methylated sites facilitating the recruitment of Dnmt1 to this strategic position. In this context, our project was aimed to further understand the mechanism of interaction of the SRA domain with hemi-methylated DNA. To this end, oligonucleotide duplexes were labeled by 2-aminopurine, a fluorescent nucleoside analogue sensitive to environment, at various positions close to the single hemi-methylated CpG recognition site. Steady-state and time-resolved fluorescence spectroscopy measurements of these duplexes bound to the SRA domain enabled us to site-specifically characterize the conformational changes induced by the binding of this domain. In agreement with the tridimensional structure of the SRA/DNA complex, our data suggest that the SRA domain is able to flip the methylcytosine while preserving the structure of the surrounding bases in the duplex. The SRA domain was shown to bind with the same mechanism to hemi-methylated, fully-methylated and non-methylated duplexes. Our data suggest the UHRF1 protein plays a role of "reader" that scans the DNA sequence for hemi-methylated sites.

Keywords : epigenetic ; DNA methylation ; UHRF1 protein ; SRA domain ; 2-aminopurine ; fluorescence spectroscopy