



HAL
open science

Étude systémique des cibles génomiques de la methyl-CpG binding domain protein 2 (MBD2), un répresseur transcriptionnel dépendant de la méthylation de l'ADN : évolution de la distribution de MBD2 dans un modèle syngénique de progression tumorale mammaire

Laury Perriaud

► **To cite this version:**

Laury Perriaud. Étude systémique des cibles génomiques de la methyl-CpG binding domain protein 2 (MBD2), un répresseur transcriptionnel dépendant de la méthylation de l'ADN : évolution de la distribution de MBD2 dans un modèle syngénique de progression tumorale mammaire. Médecine humaine et pathologie. Université Claude Bernard - Lyon I, 2010. Français. NNT : 2010LYO10216 . tel-00833153

HAL Id: tel-00833153

<https://theses.hal.science/tel-00833153>

Submitted on 12 Jun 2013

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

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

THESE DE L'UNIVERSITE DE LYON

Délivrée par

L'UNIVERSITE CLAUDE BERNARD LYON 1

ECOLE DOCTORALE BMIC

DIPLOME DE DOCTORAT

(Arrêté du 7 août 2006)

Présentée et soutenue publiquement le 3 novembre 2010

par

Melle PERRIAUD Laury

Etude systémique des cibles génomiques de la methyl-CpG binding domain protein 2 (MBD2), un répresseur transcriptionnel dépendant de la méthylation de l'ADN.

Evolution de la distribution de MBD2 dans un modèle syngénique de progression tumorale mammaire.

Directeur de thèse : Dr Robert DANTE

JURY :

Mr Joel Lachuer
Mme Frédérique Magdinier
Mr Philippe Boucher
Mr Jean Benhattar
Mr Zdenko Herceg
Mr Robert Dante

Président
Rapporteur
Rapporteur
Examineur
Examineur
Directeur de thèse

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université

Vice-président du Conseil Scientifique

Vice-président du Conseil d'Administration

Vice-président du Conseil des Etudes et de la Vie Universitaire

Secrétaire Général

M. le Professeur L. Collet

M. le Professeur J-F. Mornex

M. le Professeur G. Annat

M. le Professeur D. Simon

M. G. Gay

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard

Faculté de Médecine Lyon Sud – Charles Mérieux

UFR d'Odontologie

Institut des Sciences Pharmaceutiques et Biologiques

Institut des Sciences et Techniques de Réadaptation

Département de Biologie Humaine

Directeur : M. le Professeur J. Etienne

Directeur : M. le Professeur F-N. Gilly

Directeur : M. le Professeur D. Bourgeois

Directeur : M. le Professeur F. Locher

Directeur : M. le Professeur Y. Matillon

Directeur : M. le Professeur P. Farge

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies

Département Biologie

Département Chimie Biochimie

Département GEP

Département Informatique

Département Mathématiques

Département Mécanique

Département Physique

Département Sciences de la Terre

UFR Sciences et Techniques des Activités Physiques et Sportives

Observatoire de Lyon

Ecole Polytechnique Universitaire de Lyon 1

Institut Universitaire de Technologie de Lyon 1

Institut de Science Financière et d'Assurance

Institut Universitaire de Formation des Maîtres

Directeur : M. le Professeur F. Gieres

Directeur : M. le Professeur C. Gautier

Directeur : Mme le Professeur H. Parrot

Directeur : M. N. Siauve

Directeur : M. le Professeur S. Akkouche

Directeur : M. le Professeur A. Goldman

Directeur : M. le Professeur H. Ben Hadid

Directeur : Mme S. Fleck

Directeur : M. le Professeur P. Hantzpergue

Directeur : M. C. Collignon

Directeur : M. B. Guiderdoni

Directeur : M. le Professeur J. Lieto

Directeur : M. le Professeur C. Coulet

Directeur : M. le Professeur J-C. Augros

Directeur : M R. Bernard

RESUME

Les protéines à « Methyl-CpG-binding domain » (MBD) jouent un rôle important dans l'interprétation de la méthylation de l'ADN conduisant à la répression transcriptionnelle via le recrutement de complexes remodelant la chromatine. Dans les cancers, MBD2 jouerait un rôle essentiel dans la perte d'expression des gènes hyperméthylés. Ainsi, MBD2 serait une cible potentielle pour rétablir, en partie au moins, leur expression. Caractériser, à l'échelle du génome, la distribution de MBD2 et ses conséquences sur la répression transcriptionnelle au cours de la cancérogenèse est donc une étape incontournable. (1) L'impact sur l'expression génique de l'inhibition de MBD2 par interférence à l'ARN, a été étudié en utilisant des puces, dans des cellules normales MRC5. La perte de MBD2 n'induit pas de surexpression génique globale et la densité en CpG des promoteurs méthylés semble être une composante importante dans la force de répression par MBD2. (2) Les profils de méthylation de l'ADN, de liaisons de MBD2 et de l'ARN polymérase II dans les cellules HeLa ont été analysés par ChIP-on-chip avec des puces promoteurs. Ces mêmes approches couplées à l'analyse de l'acétylation des histones H3 ont été réalisées dans un modèle cellulaire syngénique de progression tumorale mammaire humaine. Dans les modèles étudiés, une forte proportion de gènes silencieux et méthylés est liée par MBD2. Les comparaisons entre cellules immortalisées et transformées ne montrent pas de changements majeurs de la méthylation de l'ADN ou de la répression transcriptionnelle, par contre une redistribution de MBD2 parmi ces sites est observée, suggérant une redondance entre les protéines liant l'ADN méthylé.

The Methyl-CpG Binding Domain protein 2 (MBD2), a DNA methylation-dependent transcriptional repressor: identification and characterization of MBD2 targets by genome-wide approach

ABSTRACT

The Methyl-CpG-Binding Domain (MBD) proteins represent key molecules in the interpretation of DNA methylation signals leading to gene silencing through recruitment of chromatin remodeling complexes. In cancer, a member of this protein family, MBD2, seems to play an important role in the loss of expression of aberrantly methylated genes. Thus, MBD2 may be a potential target to reestablish their expression. Mapping of MBD2 binding sites and the relationship between MBD2 binding and transcriptional activity was, therefore, a crucial step. (1) We investigated the impact of MBD2 inhibition by RNA interference on gene expression, using microarray analysis, in a normal human fibroblastic cell line, MRC-5. MBD2 depletion did not induce global gene overexpression and CpG density of the methylated promoters seems to be an important parameter in the strength of the transcriptional repression mediated by MBD2. (2) Global profiling for different layers of epigenetic modifications (DNA methylation, MBD2 association) and RNA polymerase II binding sites in HeLa cells was analyzed by a ChIP-chip method using human promoter arrays. This approach, combined with an analysis of H3 histone acetylation patterns, was performed in a syngenic model of breast cancer progression. In the models analyzed MBD2 appeared to be a true methylation-dependent transcriptional repressor. Furthermore, MBD2 binds to a high proportion of methylated silent genes. Comparisons between immortalized and transformed cells did not indicate major changes of DNA methylation or gene silencing, while a redistribution of MBD2 among these sites was observed, suggesting a redundancy between methylated binding proteins.

DISCIPLINE : Epigénétique et cancer

MOTS-CLES : Epigenetic, DNA methylation, Methyl-CpG-binding domain protein 2 (MBD2), transcriptional silencing, chromatin modifications, tumoral progression, microarrays

INTITULE ET ADRESSE DU LABORATOIRE :

Laboratoire d'Oncogénèse et Progression tumorale – Unité mixte INSERM U590
Centre Léon Bérard- 28 rue Laennec 69373 Lyon cedex 08 France

Remerciements

Je remercie le **Professeur Alain Puisieux** pour m'avoir accueillie dès mon Master au sein de son unité INSERM U590 et le **Professeur Charles Dumontet** dans son équipe « Pharmacogénétique et résistance ». Merci de m'avoir donné une structure idéale pour la réalisation de ce projet.

Bien évidemment, je remercie le **Dr Robert Dante** pour m'avoir accueillie au sein de son groupe et ainsi avoir eu l'opportunité de participer à ses projets. Merci à vous pour m'avoir fait partagé votre passion pour l'épigénétique au cours de nos nombreuses discussions. Merci de votre confiance lors de la conduite de mon projet et lors de l'encadrement d'étudiants. Merci pour l'autonomie que vous m'avez donnée. Merci également pour m'avoir fait participer à divers congrès dès mon arrivée et surtout à celui en Espagne où j'ai eu l'occasion d'exposer nos derniers travaux!

Je remercie aussi les autres membres de notre groupe.

Merci au **Pr Lucien Frappart** pour ses exposés et discussions anatomo-pathologiques sur les cancers du sein, du col de l'utérus. Merci également pour vos conseils et votre gentillesse !

Un grand merci à toutes les "Bob's girls" ! A **Amandine**, qui dès mon arrivée, a su me conseiller avec une grande efficacité ! Ton aide a été pour moi très précieuse, que ce soit pour la formation aux techniques de notre groupe que pour les corrections et suggestions pour mes rapports ! Merci également pour toutes nos discussions épigénétiques et bien moins scientifiques ! J'espère que tu obtiendras rapidement le poste que tu mérites !

Merci à **Houda**. Ton année passée avec nous a permis de nous initier au concept de récepteurs à dépendance. Merci pour ta perpétuelle bonne humeur ! Je souhaite que ta thèse à Lyon- Strasbourg se passe pour le mieux !

Merci à **Benoit** et **Rémi**, présents lors de leurs stages masters M1 et M2, d'avoir contribué à la mise au point de la précipitation d'ADN méthylé, à **Magali** alias « Schtroumpfette » pour son aide lors de ChIP.

Je remercie également nos collaborateurs. Merci aux membres de la plateforme ProfileXpert, tout particulièrement **Pr Joel Lachuer**, **Dr Catherine Rey**, **Dr Anne**

Wierinckx, Dr Cyril Lamigeon et **Séverine Croze** pour leur expertise sur les puces d'expressions (analyses et validations). Merci à Joel, Séverine et **Nicolas Nazaret** pour s'être lancés avec nous dans l'aventure « ChIP-on-chip » ! Séverine, par tes conseils techniques et ton efficacité pour les hybridations sur puces, tu as su rapidement faire évoluer les résultats. Nicolas, tes compétences en bioinformatique nous ont été très profitables ! Merci Joel d'avoir rendu cette collaboration possible, d'avoir engagé tout le personnel nécessaire de la plateforme dans ce projet afin de le mener rapidement à bien, d'avoir cru en nous pour la partie ChIP. Merci également pour toutes les discussions relatives à nos projets. J'en profite également pour te remercier d'avoir accepté d'être membre de mon jury de thèse et d'en avoir accepté la présidence.

Merci au **Dr Jean Benhattar** et à son équipe. La collaboration sur le projet « *hTERT* » s'est finie de la meilleure façon qui soit, par la publication de nos travaux. Merci également pour toutes nos interactions annuelles ou bi-annuelles sur vos et nos projets. Ces discussions nous ont permis de faire le point sur nos travaux et de les aiguiller. Je te remercie également d'avoir accepté de suivre mon travail lors de mon comité de thèse annuel. A ces occasions, tes conseils et encouragements m'ont permis d'avancer avec confiance ! Et merci d'avoir accepté d'être examinateur de ma soutenance de thèse, représentant d'une part une continuité de mon comité et marquant d'autre part l'intérêt que tu portes à mes projets.

Et merci au **Dr Anne-Pierre Morel** pour m'avoir fourni le modèle de progression tumorale mammaire.

Je remercie tous les autres membres de mon jury de thèse. Merci au **Dr Frédérique Magdinier** et **Pr Philippe Boucher** d'avoir accepté et pris le temps de juger et corriger mon manuscrit. Merci au **Dr Zdenko Herceg** d'avoir accepté d'être examinateur lors de ma soutenance de thèse.

Un merci particulier au Dr Claude Caron de Fromental d'avoir également fait partie de mon comité de thèse. Merci pour ton regard sur mes travaux et tes suggestions !

Je remercie également le CLARA et l'association ARC pour m'avoir financé et avoir cru en mon projet.

Enfin, je voudrai remercier tous mes collègues « anciennement Cheney A » aux 4^e et 5^e étages, notamment **Erika, Elodie, Kamel, Gaetan, Stéphane, Alexia**, « les 2 **Marion** »

qui ont permis que cette thèse s'effectue dans un cadre agréable durant mon M2 et début de thèse. Je remercie tous ceux « nouvellement du 2^e étage du Cheney D », avec un clin d'œil particulier pour **Freddy, Julie, Mélanie, Bastien, Agnès, Stéphanie** et la « **Dadoo's team** » avec **David, Clothilde, Hélène** et **Arnaud**. Merci surtout à **Clémence** et **Delphine** qui ont toujours répondu présentes, que ce soit sur le plan professionnel que personnel. Merci également pour votre gaieté, toutes vos blagues et toutes les sorties que l'on a pu faire et que l'on continuera de faire bien après mon départ! Clémence, je suis sûre que tu finiras ta thèse avec brio ! Et Delphine, accroches-toi, je suis persuadée qu'une belle publication est à venir ! Merci aussi à **Benjamin**, tu as toujours été à mon écoute, prêt à apporter ton aide et de façon inconditionnelle ton soutien. Merci également pour tous les fous rires à midi ou le soir, toutes tes « private jokes » vont me manquer ! Merci à **Gaël** et **Kader** d'être encore présents pour me conseiller.

Enfin, je remercie toute **ma famille** et **mes amis** qui m'ont épaulé dans cette aventure même si pour la plupart le « monde de la biologie » est un monde à part ! Un merci tout particulier à **mes parents** qui ont toujours soutenu mes projets et m'ont laissé l'indépendance nécessaire. Et promis, le chapitre « études » est clos !

Je remercie également **Céline** et **Marie** pour m'avoir permis d'intégrer l'association « science sans blues » et de m'avoir offert l'opportunité de partager ma passion pour le monde scientifique à de petits bouts de choux !

Avant-propos.....	1
Chapitre 1 : La méthylation de l'ADN, un signal épigénétique.....	3
I- La méthylation de l'ADN, un phénomène universel ou presque	3
I.A Identification de la « méthyl-cytosine »	3
I.B La méthylation de l'ADN diffère entre les espèces	3
II- Méthylation de l'ADN chez les vertébrés	4
II.A Localisation dans les cellules somatiques saines	4
II.B Définitions d'un îlot de CpG	5
II.C Rôles de la méthylation de l'ADN	6
III- Mise en place et maintien de cette marque épigénétique : rôle des méthyltransférases d'ADN.....	6
III.A Les DNMT : des gardiens nombreux pour garantir une méthylation de l'ADN.....	6
III.B Les DNMT présentes chez l'Homme.....	7
III.B.1 Descriptions génétiques et structurales des DNMT	7
III.B.1.a DNMT1	7
III.B.1.b La famille DNMT3.....	8
III.B.2 Les DNMT présentent une expression ubiquitaire.....	9
III.B.3 Les DNMT catalysent la méthylation de l'ADN	9
IV- Une marque épigénétique réversible	11
IV.A Déprogrammation et reprogrammation de la méthylation de l'ADN durant l'embryogenèse	11
IV.B Altérations des profils de méthylation dans les cancers.....	12
IV.B.1 Hypométhylation globale du génome.....	13
IV.B.2 Hyperméthylation locale des promoteurs riches en CpG.....	14
IV.B.3 Implication des DNMT dans ces altérations de profils	15
IV.B.3.a Dérégulations de l'expression des DNMT dans les cancers.....	15
IV.B.3.a.i Mise en évidence de ces dérégulations.....	15
IV.B.3.a.ii Conséquences fonctionnelles de ces dérégulations	16
IV.B.3.b Influence de l'environnement sur l'activité des DNMT	17
IV.C Conclusion.....	18
Chapitre 2 : Une lecture par les protéines liant l'ADN méthylé.....	19
I- La famille à domaine SRA	19
II- La famille des protéines à methyl-CpG-binding domain (MBD)	21
II.A Caractérisation des domaines peptidiques des protéines à MBD.....	21

II.A.1	Un domaine de liaison à l'ADN méthylé (ou MBD)	21
II.A.2	Le domaine de répression transcriptionnelle (TRD)	21
II.A.3	Les autres domaines des protéines à MBD	22
II.B	Localisations nucléaire et tissulaire des protéines à MBD	22
II.B.1	Une localisation nucléaire qui coïncide avec la présence d'ADN méthylé	22
II.B.2	Une expression tissulaire quasiment ubiquitaire	23
II.C	Une famille impliquée dans la répression transcriptionnelle	24
II.C.1	Des répresseurs transcriptionnels dépendants de la méthylation	24
II.C.1.a	Une répression dépendante de la méthylation associée à une compaction chromatinienne	25
II.C.1.a.i	Les protéines à MBD initient une répression dépendante de la méthylation de l'ADN en recrutant des désacétylases d'histones	25
	• MeCP2, premier répresseur de la famille identifié comme recruteur de HDAC	25
	• MBD2 a une fonction répressive très HDAC-dépendante	26
	• MBD1 et MBD4 jouent un faible rôle dans la répression transcriptionnelle dépendante des HDAC	27
II.C.1.a.ii	Les protéines à MBD initient une répression dépendante de la méthylation de l'ADN en recrutant des méthyltransférases d'histones	27
	• MeCP2, coREST et HP1 associé à la méthyltransférase SUV39H1	27
	• MBD1 peut interagir avec différentes HMT	28
II.C.1.a.iii	De multiples modifications de la structure chromatinienne interviennent dans la répression dépendante de la méthylation de l'ADN et initiée par les protéines à MBD	29
II.C.1.b	La protéine MeCP2 initie d'autres mécanismes dépendant de la méthylation pour remodeler la chromatine	30
II.C.1.b.i	MeCP2 et le complexe SWI/SNF remodelant la chromatine	30
II.C.1.b.ii	MeCP2 peut former des boucles d'ADN	31
II.C.1.b.iii	MeCP2 peut séquestrer des facteurs appartenant à la machinerie transcriptionnelle	31
II.C.1.c	La répression dépendante de la méthylation serait favorisée au sein de certaines séquences consensus	32
II.C.2	Des répresseurs également recrutés par des facteurs de ciblage liés à l'ADN	33
II.C.2.a	Quand le complexe Mi2/NuRD est recruté via MBD2	33
II.C.2.b	MeCP2 et PU.1	34
II.C.2.c	MBD2, MeCP2 et ARN interférants	34
II.C.3	Une activité répressive modulée	35
II.C.3.a	MBD2 et MeCP2 recrutent différents partenaires pour intervenir directement dans l'activation transcriptionnelle	35
II.C.3.b	MBD2 et MeCP2 ont aussi des rôles indirects dans l'activation transcriptionnelle	36
II.C.3.b.i	MBD2 entre en compétition avec des activateurs transcriptionnels	36
	• Compétitions de liaison à MBD2 entre HDAC1 et FAK	37
	• Compétitions de fixation à l'ADN	37
II.C.3.b.ii	MeCP2 intervient dans l'épissage alternatif	38
II.C.3.c	Les modifications post-traductionnelles des protéines à MBD modulent leur activité répressive	38
II.C.3.c.i	Phosphorylation/déphosphorylation de la protéine MeCP2	38
II.C.3.c.ii	Méthylation des arginines de la protéine MBD2	39
II.C.3.c.iii	SUMOylation de la protéine MBD1	39

II.C.4	Les protéines à MBD servent d'intermédiaires entre la méthylation de l'ADN et les changements chromatiniens conduisant à la répression transcriptionnelle	40
II.D	Balance entre spécificités de cibles et tissulaires et redondances	41
II.D.1	Spécificités d'un locus pour une protéine à MBD	56
II.D.1.a	Mises en évidence de spécificités de cibles et tissulaires	41
II.D.1.b	Comment expliquer les spécificités de cibles des protéines à MBD	42
II.D.2	Quand plusieurs protéines à MBD occupent le même site	43
II.D.3	Spécificités de cibles et tissulaires se mêlent à un phénomène de redondance	43
II.E	Implications des protéines à MBD dans d'autres processus biologiques	44
II.E.1	MeCP2, MBD2, MBD1 et leurs implications dans le développement du système nerveux	44
II.E.2	MeCP2, MBD4, MBD1 et maintien de l'intégrité du génome	45
II.E.3	Transmission des marques épigénétiques	46
II.E.4	Une famille de protéines à MBD qui couvre de nombreux mécanismes cellulaires	46
II.F	La protéine MBD2 et la cancérogenèse	47
II.F.1	Une combinaison de dérégulations du message et du lecteur	47
II.F.1.a	La protéine MBD2 présente une expression altérée dans les cancers	47
II.F.1.b	Les gènes suppresseurs de tumeurs, par l'hyperméthylation de leurs promoteurs, deviennent des cibles privilégiées de la protéine MBD2	48
II.F.2	Rôles de la protéine MBD2 dans différents cancers	49
II.F.3	La protéine MBD2, nouvelle cible pour un traitement anti-cancéreux ?	50
III-	La famille Kaiso	51
III.A	Caractérisation des domaines protéiques de la famille Kaiso	51
III.A.1	Un domaine en doigts de zinc et jusqu'à deux fonctions	51
III.A.1.a	Un domaine en doigts de zinc conférant une double spécificité de liaison à l'ADN	51
III.A.1.a.i	Identification du domaine de liaison à l'ADN	51
III.A.1.a.ii	Une liaison à l'ADN bimodale	52
III.A.1.b	Un domaine en doigts de zinc pour une interaction protéine-protéine concernant ZBTB4 et ZBTB38	53
III.A.2	Les autres domaines protéiques	53
III.A.2.a	Les autres doigts de zinc : quelles fonctions ?	53
III.A.2.b	Le domaine BTB/POZ	53
III.B	Localisations cellulaires et tissulaires des membres de la famille Kaiso	54
III.B.1	Localisation au sein des compartiments cellulaires	54
III.B.2	Localisation au sein des lignées cellulaires et des tissus	55
III.C	Rôles des protéines de la famille Kaiso dans la régulation génique	56
III.C.1	Une dualité fonctionnelle pour Kaiso et ZBTB38	56
III.C.1.a	Kaiso et ZBTB38, des répresseurs transcriptionnels	56
III.C.1.a.i	Le répresseur Kaiso se lie sur des séquences nucléotidiques via son domaine à doigts de zinc	56
•	indépendamment de la méthylation de l'ADN	56
•	après reconnaissance de CpG méthylés	57
III.C.1.a.ii	Le répresseur Kaiso se fixe à des protéines via son domaine BTB/POZ	58
III.C.1.a.iii	CIBZ/ZBTB38 et le co-répresseur CtBP	58
III.C.1.b	Kaiso et ZBTB38, également des activateurs transcriptionnels	59
III.C.2	ZBTB4 et son activité répressive	59

III.D La famille Kaiso et son implication dans les grandes fonctions biologiques chez les vertébrés.....	60
III.D.1 Kaiso, ZBTB38 et le développement embryonnaire.....	61
III.D.2 ZBTB4, ZBTB38 et l'apoptose.....	61
III.E Kaiso et la cancérogenèse.....	61
III.F Redondance fonctionnelle entre protéines liant l'ADN méthylé?.....	62
III.F.1 Redondance entre les différents membres de la famille Kaiso?.....	62
III.F.2 Redondance entre les différentes familles de protéines liant l'ADN méthylé?.....	63
III.F.2.a Quelques pistes en faveur de cette hypothèse.....	63
III.F.2.b MeCP2, Kaiso et le syndrome de Rett.....	64
III.F.2.c Conclusion: la lecture de la méthylation de l'ADN est un mécanisme complexe.....	65

Chapitre 3 : Une traduction chromatinienne pour une répression transcriptionnelle.....66

I- Les histones, protéines structurales de la chromatine.....	66
I.A Le nucléosome : unité fondamentale de la chromatine.....	66
I.B Le corps nucléoprotéique : unité fondamentale du nucléosome.....	66
II- Rôle prépondérant des histones dans la stabilisation de la chromatine.....	67
II.A Des modifications post-traductionnelles des histones associées à la régulation génique.....	68
II.B L'acétylation et la désacétylation des histones.....	68
II.B.1 Identification du groupement acétyle sur les lysines des histones.....	68
II.B.2 Le jeu des HAT et HDAC.....	69
II.B.2.a L'acétylation, une modification catalysée par des HAT.....	69
II.B.2.b Les HDAC catalysent l'excision des groupements acétyles.....	70
II.B.3 Désacétylation/acétylation des histones : un phénomène dynamique impliqué dans la régulation transcriptionnelle.....	71
II.B.3.a L'acétylation : une marque d'activation transcriptionnelle.....	71
II.B.3.a.i Corrélation entre acétylation et synthèse d'ARN.....	71
II.B.3.a.ii Corrélation entre acétylation et activation transcriptionnelle.....	71
II.B.3.b La désacétylation, une marque de répression transcriptionnelle.....	72
II.B.3.b.i Recrutement des HDAC par les protéines liant l'ADN méthylé.....	72
II.B.3.b.ii Quand les DNMT s'en mêlent.....	73
II.B.3.b.iii Corrélation entre désacétylation et répression transcriptionnelle.....	73
II.B.3.c HDAC et HAT : rôles spécifiques et complémentaires.....	74
II.B.3.c.i HDAC et HAT pour une même lysine?.....	74
II.B.3.c.ii Coopérations, compétitions ou régulations entre HDAC et HAT.....	74
II.C La méthylation des histones.....	75
II.C.1 Une modification sur des arginines et lysines.....	75
II.C.2 Une modification catalysée par des HMT.....	75
II.C.3 La méthylation des histones, une marque associée à l'activation ou à la répression transcriptionnelle.....	76
II.C.4 Les déméthylases.....	77

II.D	Combinaison de méthylation et d'acétylation sur les histones	78
------	--	----

Chapitre 4 : Des technologies pour des études à grande échelle du signal de méthylation de l'ADN, de ses lecteurs et des changements chromatiniens associés.....80

I-	Cartographies des sites méthylés de l'ADN.....	80
I.A	A partir de digestions enzymatiques	80
I.B	A partir de modifications chimiques par bisulfite.....	81
I.B.1	A partir de précipitations d'ADN ou de chromatine	82
I.B.2	A partir d'une combinaison de techniques.....	82
II-	Cartographies des sites de fixation des protéines.....	83
II.A	Modification d'ADN par DAM fusionnée.....	83
II.B	Immunoprécipitation de chromatine avec des anticorps.....	83
III-	Une cartographie des modifications des histones.....	84
IV-	Un choix de puces à ADN déterminant.....	84

Objectifs de travail.....85

Résultats.....87

Article 2:	Relationship Between the CpG content of Gene Promoter and Transcriptional Repression by the Methyl-CpG Binding Domain Protein 2 in Normal Human Fibroblastic Cells.....	87
-------------------	---	----

Article 3:	Methyl-CpG Binding Domain Protein, a DNA Methylation-Dependent Gene silencer, Genome-Wide Analysis.....	123
-------------------	---	-----

Article 4:	Specific Distribution of the Methyl-CpG Binding Domain Protein 2 at methylated DNA sequence during tumoral progression, in a syngenic model of breast cancer.....	178
-------------------	---	-----

Discussion.....202

- 1. Mise au point d'un outil permettant la sélection spécifique de régions d'ADN méthylées quelque soit leur densité en doublets CpG.....202
- 2. La protéine MBD2, un répresseur transcriptionnel dépendant de la méthylation de l'ADN.....203
 - 2.1 Caractéristiques des sites fixés par la protéine MBD2.....203
 - 2.1.1 La protéine MBD2 lie spécifiquement des régions méthylées.....203
 - 2.1.2 La protéine MBD2 se concentre dans les régions comprises entre -700 et +200pb autour du site d'initiation de la transcription.....204
 - 2.1.3 La liaison de MBD2 ne dépend pas de la densité en CpG.....205
 - 2.1.4 La protéine MBD2 ne se lie à aucune séquence consensus.....206
 - 2.2 La protéine MBD2 joue un rôle majeur dans la répression transcriptionnelle.....206
 - 2.3 La répression dépendante de la méthylation initiée par la protéine MBD2 est associée à une compaction chromatinienne.....207
- 3. La force de répression de la protéine MBD2 est modulable.....208
 - 3.1 La force de répression de MBD2 dépend de la distance au TSS.....208
 - 3.2 La force de répression de MBD2 dépend de la densité en CpG.....209
- 4. La protéine MBD2 est redistribuée au cours de la progression tumorale.....210
- 5. La protéine MBD2, une piste séduisante pour un traitement anti-cancéreux.....212
 - 5.1 La répression de la protéine MBD2 est réversible.....212
 - 5.2 La perte de la protéine MBD2 aurait un impact limité sur la régulation transcriptionnelle.....212
 - 5.2.1 Un faible nombre de gènes touchés par l'absence de MBD2.....212
 - 5.2.2 Redondance fonctionnelle entre MBD2 et d'autres protéines liant l'ADN méthylé ?.....213

Conclusions et perspectives.....215

Références bibliographiques.....218

Annexes.....246

Annexe 1: article 1: Specific association between the methyl-CpG-binding domain
protein 2 and the hypermethylated region of the human telomerase reverse
transcriptase promoter in cancer cells.....246

Annexe 2 : publications.....253

: communications affichées.....254

Table des illustrations

Figure 1 : Catalyse de la méthylation de l'ADN sur le carbone 5 d'une cytosine par les méthyltransférases d'ADN (DNMT).....Verso page 2

Figure 2 : Corrélations entre méthylation de l'ADN et répression transcriptionnelle en fonction de la densité en CpG des régions promotrices.....Verso page 5

Figure 3 : Inhibition directe de la fixation d'un facteur de transcription (FT) sur sa séquence cible lorsqu'elle devient méthylée.....Verso page 5

Figure 4 : Mécanisme de méthylation *de novo* par DNMT1, suite au recrutement du complexe DNMT1-Np95 au niveau des foyers de réplication.....Verso page 8

Figure 5 : Modifications des profils de méthylation au cours du développement embryonnaire.....Verso page 10

Figure 6 : Altérations des profils de méthylation au cours de la cancérogenèse...Verso page 12

Figure 7 : Influence de l'environnement sur l'activité méthyltransférase d'ADN.....Verso page 16

Figure 8 : Schématisation protéique des membres fondateurs des deux familles de protéines à domaine SRA.....Verso page 18

Figure 9 : Schématisation protéique de la famille des protéines à « methyl-CpG binding domain » (protéines à MBD).....Verso page 20

Figure 10 : Les protéines à MBD font le lien entre méthylation de l'ADN, répression transcriptionnelle et compaction chromatinienne.....Verso page 24

Figure 11 : Le répresseur MBD2 agit en recrutant des complexes contenant des HDAC suite à sa liaison sur l'ADN méthylé.....Verso page 25

<u>Figure 12</u> : Le répresseur MBD1 recrute des complexes contenant des HDAC et/ou des HMT lorsqu'il est lié à l'ADN méthylé.....	Verso page 26
<u>Figure 13</u> : Répression transcriptionnelle dépendante de la méthylation de l'ADN initiée par la protéine MeCP2.....	Verso page 29
<u>Figure 14</u> : Régulation transcriptionnelle de la myogénine par compétition de liaison entre HDAC1 et FAK phosphorylée pour la protéine MBD2.....	Verso page 36
<u>Tableau 1</u> : Gènes hyperméthylés dans les cancers et ciblés par la protéine MBD2.....	Verso page 47
<u>Figure 15</u> : Schématisation protéique de la famille Kaiso.....	Verso page 50
<u>Figure 16</u> : Motifs des séquences consensus pour la liaison de Kaiso (KBS) et de ZBTB4 (Z4BS).....	Verso page 51
<u>Figure 17</u> : Répressions transcriptionnelles dépendantes de la méthylation de l'ADN initiées par la protéine Kaiso.....	Verso page 56
<u>Figure 18</u> : Seul cas de répression transcriptionnelle rapporté pour CIBZ.....	Verso page 57
<u>Figure 19</u> : ZBTB4 est un répresseur transcriptionnel de <i>p21Cip21</i>	Verso page 59
<u>Figure 20</u> : Formation d'un nucléosome, l'unité fondamentale de la chromatine.....	Verso page 65
<u>Figure 21</u> : La méthylation et l'acétylation, deux modifications post-traductionnelles associées à la régulation transcriptionnelle.....	Verso page 67
<u>Figure 22</u> : Interactivités des marques d'acétylation et de méthylation selon leurs positions sur les histones H3 et H4.....	Verso page 77

Liste des principales abréviations

ADN : Acide desoxyribonucléique complémentaire de l'ARN messager

ARNi : interférence à l'ARN

ARNm : acide ribonucléique messager

BTB/POZ : Broad Complex, Tramtrak, Bric à brac/ Pox virus and zing finger

ChIP : *chromatin immunoprecipitation*, immunoprécipitation de chromatine

ChIP-on-chip : Hybridation sur puce (chip) de fractions génomiques immunoprécipitées par ChIP

CpG : cytosine et guanine adjacentes reliées par un phosphate

DNMT : *DNA methyltransferase*, méthyltransférases d'ADN

H3/H4 : histone 3/ histone 4

H3K9 : lysine 9 de l'histone 3

HAT : *histone acetyltransferase*, acétyltransférase d'histones

HDAC : *histone deacetylase*, désacétylases d'histones

HMT : *histone methyltransferase*, méthyltransférases des histones

hTERT : *human telomerase reverse transcriptase*, transcriptase inverse de la télomérase humaine

kb : kilobases

kDa : kiloDaltons

MBD : « *Methyl-CpG-Binding Domain* », domaine de liaison à l'ADN méthylé

MeDP : *Methylated DNA precipitation*, précipitation d'ADN méthylé

MeDP-on-chip : Hybridation sur puce (chip) de fractions d'ADN précipitées par la protéine recombinante 4xMBD

pb : Paires de bases

PCR : *Polymerase Chain Reaction*, réaction de polymérisation en chaîne

RT-PCR : *Reverse Transcription – Polymerase Chain Reaction*, transcription inverse suivie d'une réaction de polymérisation en chaîne

SET : Su(var)3-9, enhancer of Zeste, Trithorax

SRA : SET and RING finger associated

TRD : *transcriptional repression domain*, domaine protéique de répression transcriptionnelle

TSA : Trichostatine A

Introduction
Revue bibliographique

Avant-propos

Cette thèse s'inscrit dans le domaine de l'épigénétique. Aussi, il est apparu important de définir ce terme.

On attribue la paternité de « l'épigénétique » dans son sens moderne au biologiste Conrad H. Waddington qui en 1942 assimile ce mot à tout « phénotype héréditaire stable résultant de changements dans un chromosome sans altération dans la séquence d'ADN ».

En évoluant, la définition est devenue plus restrictive. De nos jours, l'épigénétique correspond dans son sens le plus courant à « l'ensemble des modifications héréditaires et réversibles de l'expression des gènes sans altération des séquences nucléotidiques ». Ainsi, ce terme se concentre sur les gènes et leurs régulations non expliquées par des mutations d'ADN. De plus, il y a conservation du caractère héréditaire qui implique une transmission même en absence du signal déclencheur, supposant une auto-propagation entre cellules par mitose et/ou méiose et/ou entre générations d'individus.

La méthylation de l'ADN (Jaenisch and Bird 2003) et les systèmes Polycomb et Trithorax (Bird 2007) répondent à tous les critères de cette définition. Par contre, le classement des modifications covalentes des histones comme marques épigénétiques est plus discuté (Ptashne 2007). En effet, il existe des interactions directes entre les enzymes responsables de la méthylation de l'ADN et celles de la désacétylation des histones (Fuks, Burgers et al. 2000; Bai, Ghoshal et al. 2005) et de leur méthylation (Lehnertz, Ueda et al. 2003). Aussi, certaines modifications des histones peuvent être, à titre indirect, qualifiées d'épigénétiques.

Le paramètre de l'hérédité étant problématique, il a été proposé de l'éliminer de la définition ou d'en faire un élément facultatif. Le terme « épigénétique » pourrait alors désigner une « adaptation structurelle des régions chromosomiques afin d'enregistrer, signaler ou perpétuer les altérations géniques » (Bird 2007). Cependant, une variation récente de cette définition (Berger, Kouzarides et al. 2009) comprend l'ajout des notions de signaux

déclenchement et de maintien des changements chromatiniens, replaçant l'hérédité au cœur des débats.

Ainsi, pour l'instant, le terme « épigénétique » désigne clairement la méthylation de l'ADN, les systèmes Polycomb et Trithorax, la méthylation et l'acétylation des histones.

Au cours de l'introduction de ce manuscrit, seront abordés et développés la méthylation de l'ADN et ses interprètes ainsi que les changements chromatiniens associés. Une dernière partie fera l'état des lieux des technologies actuellement disponibles pour initier une étude à grande échelle de ces mécanismes.

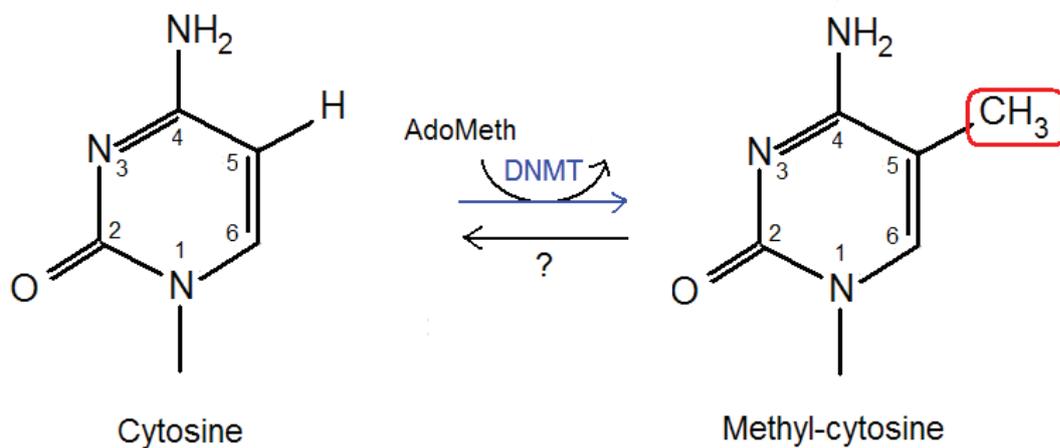


Figure 1 : Catalyse de la méthylation de l'ADN sur le carbone 5 d'une cytosine par les méthyltransférases d'ADN (DNMT)

Les DNMT réalisent le transfert d'un groupement méthyle (-CH₃) du substrat S-adenosyl-méthionine (AdoMet) sur le carbone 5 d'une cytosine, ce qui donne la « méthyl-cytosine ». Les mécanismes de déméthylation de l'ADN ne sont pas clairement établis (?). Plusieurs hypothèses sont avancées comme la substitution de la base azotée méthylée par excision via des ADN glycosylases ou des enzymes plus spécifiques, des ADN déméthylases puis remplacement par une cytosine non méthylée. Les protéines MBD2, DNMT3A et DNMT3B pourraient jouer le rôle d'ADN déméthylases même si leur implication donne encore lieu à de nombreuses controverses.

La cytosine et la méthyl-cytosine sont représentées sous leurs formules semi-développées.

La méthylation de l'ADN, un signal épigénétique

I- La méthylation de l'ADN, un phénomène universel ou presque

I.A Identification de la « méthyl-cytosine »

La méthylation de l'ADN correspond à l'addition covalente d'un groupement méthyle (-CH₃) provenant du co-facteur S-adenosyl-méthionine (AdoMet) sur une base nucléotidique, catalysée par une méthyltransférase d'ADN (DNMT).

L'identification d'un groupement méthyle sur la cytosine, une base azotée, a reposé sur l'hydrolyse acide prolongée d'ADN extrait de thymus de veau (Hotchkiss, 1948) et donné l'« épi-cytosine », renommée par la suite « méthyl-cytosine » (**Figure 1**).

I.B La méthylation de l'ADN diffère entre les espèces

La méthylation de l'ADN est un phénomène universel. Il existe néanmoins une certaine disparité parmi les êtres vivants, tant de par la nature des bases cibles de la méthylation, leur contexte de séquence, que de par leur répartition au sein des génomes.

Ainsi, la méthylation en C5 de cytosines et en N6 d'adénines est retrouvée dans le génome de nombreux champignons, procaryotes et protistes alors que la méthylation en N4 de cytosines ne concerne que les bactéries (Palmer and Marinus 1994; Cheng 1995). Chez les végétaux et les animaux, la méthylation n'est présente qu'en C5 de cytosines. Certaines exceptions sont cependant à noter : en effet, aucune méthylation de l'ADN n'est observée chez les levures *Saccharomyces cerevisiae* et *Schizosaccharomyces pombe* (champignons) ou chez le nématode *Caenorhabditis elegans* (invertébré).

La position du nucléotide méthylé au sein de l'ADN est variable entre les règnes. En effet, chez les végétaux (Finnegan, Peacock et al. 2000) et les ascomycètes *Neurospora crassa* (Selker, Tountas et al. 2003), la méthylation des cytosines peut s'effectuer au niveau de séquences symétriques 5'-CpG-3' ou 5'-CpNpG-3' avec N une base azotée quelconque, ou de séquences non symétriques 5'-CpApT-3' ou 5'-CpTpT-3'. Chez les animaux, un contexte

5'-CpG-3' permet l'ajout d'un groupement méthyle, quelques rares cas chez l'Homme dans un contexte 5'-CpA-3' ou 5'-CpT-3' dans des cellules souches embryonnaires (Ramsahoye, Biniszkiewicz et al. 2000) ou CC(A/T)GG (méthylation sur la 2^e cytosine) dans des lymphomes de cellules B matures (Malone, Miner et al. 2001) ou découvert plus récemment dans un contexte 5'-CpNpG-3' dans plusieurs types de tissus normaux ou cancéreux (Lee, Jang et al.).

D'autre part, la répartition de la méthylation est homogène pour les végétaux et les procaryotes, avec une particularité chez les bactéries puisque tous les sites méthylables sont méthylés. Par contre, la distribution de la méthylation est non uniforme pour le règne animal. En effet, les cytosines méthylées sont concentrées en de larges domaines séparés par des domaines non méthylés.

Enfin, les taux de méthylation de l'ADN sont très hétérogènes au sein du règne animal. En effet, chez l'insecte *Drosophila melanogaster*, organisme modèle en génétique possédant un petit génome, la méthylation concerne <0,1% des cytosines totales (Lyko, Ramsahoye et al. 2000) alors que chez les vertébrés ce taux peut atteindre 3 à 8%. Ces données vont dans le sens de la théorie développée par A. Bird (Bird 2002). En effet, il existerait une convergence vers une spécificité pour les cytosines, une augmentation du taux global de méthylation observée en association avec une complexification des génomes.

II- Méthylation de l'ADN chez les vertébrés

II.A Localisation dans les cellules somatiques saines

Chez les vertébrés, la méthylation concerne très majoritairement les cytosines et, dans une large proportion, celles présentes au sein de séquences symétriques 5'-CpG-3' appelées doublets CpG. Trois à 8% des cytosines totales sont méthylées (Ehrlich, Gama-Sosa et al. 1982), ce qui concerne 60 à 90% des CpG. Il existe toutefois des écarts autour de ces valeurs selon les différents tissus, comme déterminé par chromatographie en phase liquide à haute performance (HPLC) avec le thymus et le cerveau des tissus très méthylés (~1 cytosine méthylée pour 100 bases d'ADN) et le sperme et le placenta peu méthylés (~0,8) (Ehrlich, Gama-Sosa et al. 1982).

Ces dinucléotides sont répartis de façon non uniforme dans le génome. Une exclusion de la méthylation est majoritairement observée pour des séquences riches en CpG comme le montrent des expériences de digestion d'ADN murin avec l'enzyme HpaII coupant les sites non méthylés CCGG (Bird, Taggart et al. 1985). Ces séquences, appelées îlots de CpG, sont estimées à 30.000 ce qui représenterait 1% du génome. Les expériences menées par Antequera et ses collaborateurs ont ensuite porté ce nombre à 37.000 pour le génome murin et 45.000 pour le génome humain (Antequera and Bird 1993). Il existe cependant des exceptions, avec une méthylation des îlots CpG jouant un rôle dans l'inactivation du chromosome X (Mohandas, Sparkes et al. 1981) et l'empreinte parentale en éteignant un des deux allèles parentaux (Bartolomei 2009).

II.B Définitions d'un îlot de CpG

Les îlots de CpG correspondent à des séquences riches en CpG. Elles paraissent avoir échappé aux mutations de cytosines méthylées en thymine. En effet, dans le génome humain, la fréquence globale de CpG c'est-à-dire le taux global de CpG observé par rapport au taux théorique est de $\sim 1/4$ (Aerts, Thijs et al. 2004).

Les îlots de CpG constituent des séquences de minimum 200pb avec une composition en C+G plus importante que celle en A+T (ratio C+G > 50%) et une fréquence supérieure à 0,6, et retrouvées entre -500 et +1500 autour du site d'initiation de la transcription (Antequera and Bird 1993; Ioshikhes and Zhang 2000; Aerts, Thijs et al. 2004).

Ils sont généralement inclus sur la région promotrice de tous gènes de ménage et de plus de la moitié des gènes tissus-spécifiques du génome humain (Gardiner-Garden and Frommer 1987; Larsen, Gundersen et al. 1992; Bird 2002). Cette relation entre promoteurs et haute densité en CpG semble spécifique aux vertébrés car elle n'est pas retrouvée chez d'autres métazoaires comme le poisson Fugu ou la Drosophile (Aerts, Thijs et al. 2004).

Néanmoins, la classification des séquences selon le critère « îlots de CpG » ne semble aujourd'hui plus être binaire. En effet, l'évolution des techniques permettant une analyse globale du génome a permis d'établir de nouveaux seuils pour les définir, répondant mieux aux relations entre présence de méthylation, répression transcriptionnelle et changements chromatiniens déjà identifiées (Keshet, Lieman-Hurwitz et al. 1986). Les îlots de CpG sont des zones très riches en CpG (HCP) de > 500pb qui répondent à un ratio > 0,75 et une

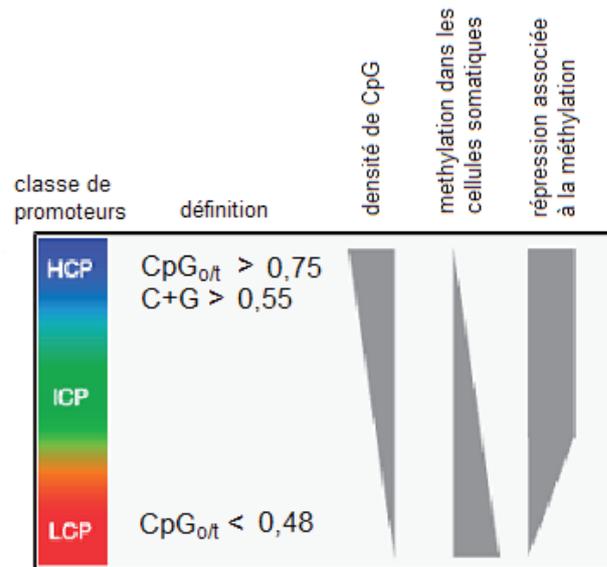


Figure 2 : Corrélations entre méthylation de l'ADN et répression transcriptionnelle en fonction de la densité en CpG des régions promotrices (Figure adaptée de (Weber, Hellmann et al., 2007))

Représentation schématique des 3 classes de promoteurs (HCP pour High-CpG promoters, LCP pour Low-CpG promoters et ICP pour intermediate-CpG promoters) définis selon la fréquence en C+G et le ratio $CpG_{observé/théorique}$ ($CpG_{o/t}$). Il existe une corrélation inverse entre la densité en CpG et le taux de méthylation de l'ADN. La répression transcriptionnelle est essentiellement associée des ICP et HCP méthylés.

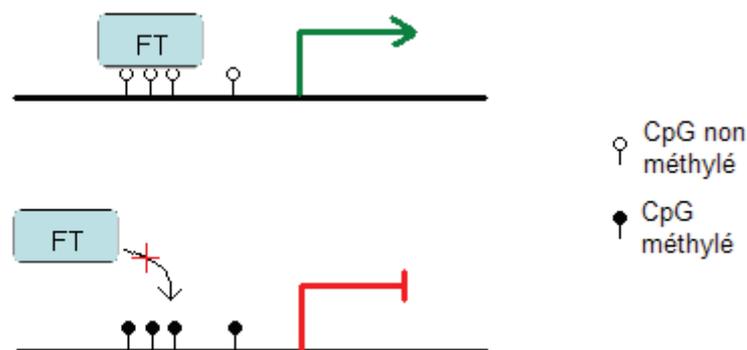


Figure 3 : Inhibition directe de la fixation d'un facteur de transcription (FT) sur sa séquence cible lorsqu'elle devient méthylée

fréquence $> 0,55$ donc plus hauts que précédemment et sont à distinguer des régions très pauvres en CpG (LCP) (ratio $< 0,48$) qui ne sont pas corrélées à une répression transcriptionnelle et des régions intermédiaires (ICP) qui ne sont ni des LCP ni des HCP (Saxonov, Berg et al. 2006; Weber, Hellmann et al. 2007) (**Figure 2**).

II.C Rôles de la méthylation de l'ADN

La méthylation de l'ADN intervient chez les vertébrés principalement dans le développement avec l'inactivation d'un des chromosomes X chez la femelle, l'empreinte parentale, la différenciation cellulaire, tous ces phénomènes étant majoritairement dû à son rôle primordial dans la régulation de l'expression génique. Elle est également impliquée dans la défense et l'intégrité structurale du génome en évitant les recombinaisons chromosomiques ou la mobilité d'éléments transposables.

La corrélation entre la méthylation de l'ADN et la répression transcriptionnelle est connue depuis 30 ans. La méthylation de l'ADN peut directement empêcher certains facteurs de transcription de se fixer sur des séquences régulatrices comme NF-kappaB mais ce mécanisme est jugé peu fréquent *in vivo* (**Figure 3**). Elle semble essentiellement intervenir dans la répression transcriptionnelle de façon indirecte en passant par la mise en place de changements chromatiniens, comme nous le verrons au cours de cette introduction (Keshet, Lieman-Hurwitz et al. 1986).

III- Mise en place et maintien de cette marque épigénétique : rôle des méthyltransférases d'ADN

III.A Les DNMT : des gardiens nombreux pour garantir une méthylation de l'ADN

La méthylation de l'ADN est indispensable chez les mammifères et les plantes alors qu'elle ne l'est pas chez *Neurospora crassa* (Selker 1998). Cette nécessité est peut être à mettre en relation avec le nombre d'ADN méthyltransferases (DNMT) retrouvées dans ces

espèces. En effet, seule une ADN méthyltransférase est présente chez *N.crassa* (DIM-2) (Kouzminova and Selker 2001) alors que les plantes peuvent en compter jusqu'à 20 (Finnegan, Peacock et al. 2000) et chez les mammifères 3 classes de DNMT ont été identifiées.

III.B Les DNMT présentes chez l'Homme

Il existe 3 classes de DNMT : DNMT1, DNMT2 et la famille DNMT3. Toutes ces protéines, à l'exception de DNMT2 (Okano, Xie et al. 1998) et DNMT3L, jouent un rôle dans l'addition de groupements méthyles à l'ADN grâce à leur activité méthyltransférase. Cet ajout permet soit de maintenir la méthylation de l'ADN c'est-à dire de reproduire le patron de méthylation après réplication de l'ADN suite à la mitose ou meiose soit d'établir un patron de méthylation propre à l'embryon et à ce moment-là elle est qualifiée de méthylation *de novo*.

III.B.1 Descriptions génétiques et structurales des DNMT

III.B.1.a DNMT1

La première DNMT identifiée a été DNMT1 chez la souris (Bestor, Laudano et al. 1988) puis chez l'Homme (Yen, Vertino et al. 1992). Ces orthologues partagent globalement 74% d'homologies avec en C-terminal certaines régions très similaires dont 8 (I, II, IV, VI, VII, VIII, IX, X) en commun avec les DNMT présentes chez les procaryotes. L'ordre de ces motifs est conservé entre espèces donc ils semblent responsables des mêmes mécanismes. Chez les procaryotes, les motifs I, II, III et VI, VIII, X formeraient respectivement des poches de liaison au co-facteur AdoMet et à la cible, le motif IV contiendrait le domaine catalytique pour le transfert de groupements méthyles alors que les motifs V et VII préserveraient le repliement du domaine de reconnaissance de la cible (Lauster, Trautner et al. 1989).

Depuis sa découverte, plusieurs fragments de séquences ont été ajoutés entre 1996 et 1997 (Yoder, Yen et al. 1996; Pradhan, Talbot et al. 1997). Des variants possédant tous une activité catalytique *in vitro* ont été décrits pour l'Homme et la souris (Pradhan, Talbot et al. 1997; Bonfils, Beaulieu et al. 2000). Des isoformes sexe-spécifiques ont même été décrit

(Mertineit, Yoder et al. 1998; Doherty, Bartolomei et al. 2002). La caractérisation d'une nouvelle région en 5' du gène *DNMT1* chez l'Homme et la souris (Yoder, Yen et al. 1996) a même permis d'orienter les recherches sur la régulation de ce gène. Ainsi, par exemple, un système de gène rapporteur, des expériences de retard sur gel (shift) et de supershift par utilisation d'anticorps ont permis de déduire que le promoteur *Dnmt1* contient un motif ciblé par Sp1 ou Sp3 pour une activation transcriptionnelle. Sp3 agirait en fait en coopération avec l'acétyltransférase d'histone p300 (Kishikawa, Murata et al. 2002) pendant la phase S/G1 (Kishikawa, Murata et al. 2003).

III.B.1.b La famille DNMT3

La famille DNMT3 compte plusieurs membres (DNMT3a, DNMT3b et DNMT3L) (Okano, Xie et al. 1998; Aapola, Kawasaki et al. 2000) partageant une similarité structurale avec DNMT1 pour la région catalytique responsable de la méthylation de l'ADN et le domaine N-terminal. Tous ces membres présentent également un domaine riche en cystéine très homologue à un domaine de la protéine ATRX contenant le motif PHD (plant homeodomain) imparfait.

Les protéines DNMT3a et 3b présentent une activité catalytique méthyltransférase contrairement à DNMT3L (DNMT3-like) qui apparaît plutôt comme un co-facteur de ces protéines pour l'établissement de l'empreinte parentale maternelle dans les lignées germinales (Bourc'his, Xu et al. 2001; Hata, Okano et al. 2002) en stimulant leur activité catalytique sans les guider sur des sites particuliers (Suetake, Shinozaki et al. 2004).

Les protéines DNMT3a et DNMT3b présentent plusieurs variants chez l'Homme et la souris (Okano, Xie et al. 1998; Robertson, Uzvolgyi et al. 1999; Chen, Ueda et al. 2002). Les deux variants de DNMT3a arborent une activité méthyltransférase *in vitro* et la forme tronquée paraît plutôt présente dans l'euchromatine et prédominante dans les tissus embryonnaires. Concernant DNMT3b, deux des cinq variants humains ont un domaine catalytique méthyltransférase tronqué donc non fonctionnel. Selon les tissus, il y a une prédominance d'un variant par rapport aux autres, comme dans le cerveau, ou une expression similaire de toutes les formes, comme dans les testicules. Aussi les rôles de DNMT3a et DNMT3b ne paraissent pas se limiter à méthyler l'ADN.

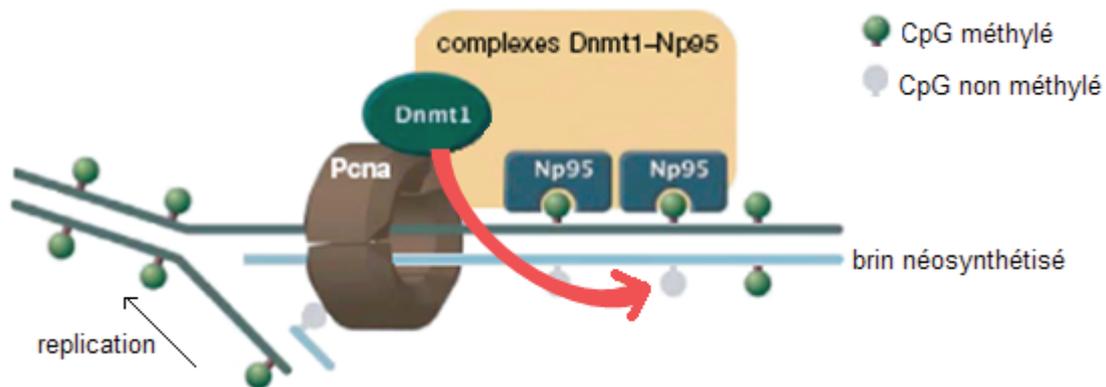


Figure 4 : Mécanisme de méthylation *de novo* par DNMT1, suite au recrutement du complexe DNMT1-Np95 par PCNA au niveau des foyers de réplication (adaptée de (Sharif, Muto et al., 2007))

Le complexe DNMT1-Np95 interagit avec PCNA via DNMT1 et avec les CpG méthylés via Np95. La protéine DNMT1 peut alors exercer son activité de méthyltransférase d'ADN sur les cytosines du brin néosynthétisé au cours de la réplication. Chaque site hémi-méthylé devient alors méthylé.

III.B.2 Les DNMT présentent une expression ubiquitaire

L'expression de *DNMT1*, *DNMT3a* et *DNMT3b* a été évaluée dans une grande série de tissus humains fœtaux et adultes normaux (Robertson, Uzvolgyi et al. 1999). Des expériences de Northern Blot ont ainsi montré que ces expressions sont ubiquitaires avec tout de même des variations entre les tissus. En effet, les taux d'ARNm *DNMT1* et *DNMT3a* les plus forts sont retrouvés dans les tissus fœtaux et *DNMT3b* semble présenter généralement un taux bien plus faible dans tous les tissus par rapport à *DNMT1* et *DNMT3a*. De plus, au vu des données, DNMT1 apparaît être la DNMT prépondérante dans les tissus adultes. Des expériences quantitatives de RT-PCR dans les tissus normaux ont permis de préciser que les distributions d'ARNm *DNMT1*, *DNMT3a* et *DNMT3b* sont assez similaires sauf concernant le muscle squelettique et les cellules mononucléaires du sang périphérique. Les taux les plus forts sont observés pour le cœur, les reins, le placenta.

III.B.3 Les DNMT catalysent la méthylation de l'ADN

Il apparaît que les DNMT peuvent intervenir dans différents processus comme la répression transcriptionnelle (Bachman, Rountree et al. 2001; Fuks, Burgers et al. 2001) mais leur rôle principal est de catalyser la méthylation de l'ADN sur les cytosines.

DNMT1 paraît intervenir préférentiellement sur des sites hémi-méthylés (Pradhan, Talbot et al. 1997; Bacolla, Pradhan et al. 1999; Pradhan, Bacolla et al. 1999), via un complexe contenant notamment la protéine NP95 (nuclear protein 95) liant l'ADN méthylé (Bostick, Kim et al. 2007) et recruté par PCNA (Proliferating Cell Nuclear Antigen) en phase S au niveau des foyers de réplication et/ou de réparation de l'ADN (Leonhardt, Page et al. 1992; Chuang, Ian et al. 1997; Baylin, Herman et al. 1998; Margot, Cardoso et al. 2001; Iida, Suetake et al. 2002; Sharif, Muto et al. 2007) (**Figure 4**). La génération par recombinaison homologue d'un variant *DNMT1* tronqué d'une région exonique contenant le site de liaison de PCNA et possédant quand même une activité catalytique *in vitro* et *in vivo* comme le montrent des expériences de FRAP a permis d'établir que l'association DNMT1-PCNA n'est pas nécessaire mais potentialise l'effet de DNMT1 pour la maintenance de la méthylation de l'ADN (Rhee, Jair et al. 2000; Spada, Haemmer et al. 2007).

Même si DNMT1 a une préférence notable pour les sites hémi-méthylés, elle est également capable de méthyler des sites initialement non méthylés (Pradhan, Talbot et al. 1997; Pradhan, Bacolla et al. 1999). Aussi, elle semble intervenir dans la méthylation de maintenance et la méthylation *de novo*.

Des mutations par recombinaison homologue de *Dnmt1* ont permis de suggérer l'existence d'autres méthyltransférases d'ADN dès le stade embryonnaire précoce (Lei, Oh et al. 1996). La famille DNMT3 a ainsi été caractérisée par la suite (Okano, Xie et al. 1998). La génération de cellules souches embryonnaires (ES) *Dnmt3a*^{-/-} ou *Dnmt3b*^{-/-} ou doublement mutées et transfectées avec un provirus non méthylé a mis en évidence une redondance de ces 2 protéines dans la méthylation *de novo in vitro*. La création de souris hétérozygotes puis homozygotes pour ces invalidations de gènes a corroboré ces résultats *in vivo* et montré que ces protéines sont indispensables aux développements embryonnaire et post-natal (Hsieh 1999; Okano, Bell et al. 1999). De plus, leur activité serait influencée par l'environnement car le recrutement de Dnmt3a par EZH2 près de sites non méthylés *in vitro* n'entraîne pas de méthylation (Rush, Appanah et al. 2009).

D'autre part, une activité de maintenance de la méthylation de l'ADN n'est plus à exclure pour les protéines DNMT3a et DNMT3b. En effet, elles ne présentent pas d'affinité particulière entre les oligonucléotides non méthylés et hémi-méthylés (Kim, Ni et al. 2002) et l'invalidation de *DNMT1*, contrairement aux résultats trouvés chez la souris (Lei, Oh et al. 1996) n'entraîne une diminution que de 20% du taux de méthylation dans des cellules humaines colorectales cancéreuses (Rhee, Jair et al. 2000). Dans ces lignées, l'invalidation de *DNMT3B* par recombinaison homologue dirigée semble peu altérer le taux de méthylation de l'ADN alors que la double invalidation de ces gènes conduit à une hypométhylation de l'ADN de l'ordre de 95% avec un impact sur tout le génome (Rhee, Bachman et al. 2002).

Enfin, si les régions C-terminales des DNMT1, DNMT3a et DNMT3b supportent l'activité catalytique de ces enzymes, les régions N-terminales permettraient de réguler leur activité. Il a même été suggéré que les DNMT1 et DNMT3a ou DNMT3b pourraient interagir via cette région afin de coopérer et stimuler l'action de DNMT1 concernant la maintenance de la méthylation de l'ADN suite à l'intervention de DNMT3a ou DNMT3b sur des sites non méthylés donc après un début de méthylation *de novo* (Kim, Ni et al. 2002).

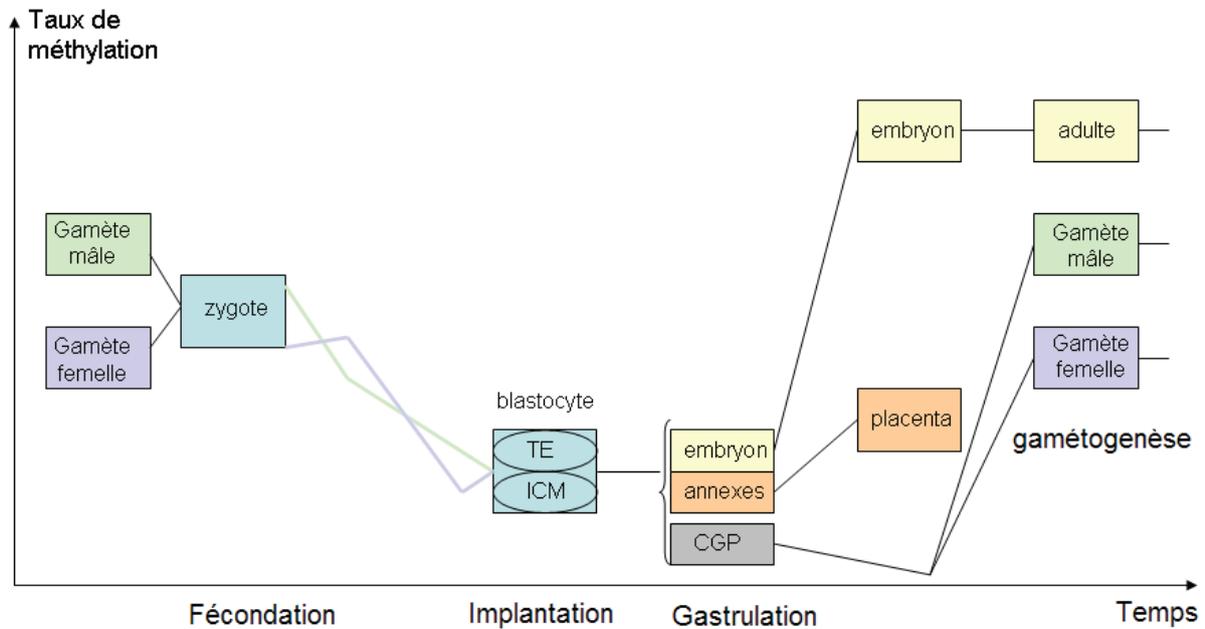


Figure 5 : Modifications des profils de méthylation au cours du développement embryonnaire

Après la fertilisation, dans le zygote, le génome paternel subit deux vagues de déméthylation tandis que le génome maternel subit une petite méthylation *de novo* avant d'être déméthylé au fur et à mesure des divisions cellulaires. Après une vague de méthylation *de novo*, le génome maternel retrouve un taux de méthylation similaire au génome paternel.

Dès le stade blastocyste, des spécialisations cellulaires ont lieu. A partir du trophoderme se développe l'embryon qui subit une vague de méthylation *de novo* pour créer son propre profil de méthylation de l'ADN. Le placenta, issu des annexes embryonnaires, reste largement sous-méthylé par rapport à l'embryon. Les cellules germinales primordiales (CGP) subissent une vague de déméthylation pour une déprogrammation complète avant d'être reprogrammé différemment selon les gamètes, avec un taux de méthylation plus important dans les gamètes mâles que femelles.

Les protéines DNMT1, DNMT3a et DNMT3b seraient donc les acteurs de la maintenance de la méthylation de l'ADN et la méthylation *de novo* avec des rôles partiellement redondants.

IV- Une marque épigénétique réversible

IV.A Déprogrammation et reprogrammation de la méthylation de l'ADN durant l'embryogenèse

Avant l'implantation de l'embryon, les génomes maternel et paternel semblent subir des changements au niveau de la méthylation d'ADN. Il a été montré chez la souris 2 vagues de déméthylation et reméthylation (**Figure 5**): la première a lieu pendant le développement cellulaire germinale et l'autre se passe après la fertilisation comme le montre l'étude de séquences répétées par utilisation d'anticorps dirigés contre la 5-méthylcytosine (Mayer, Niveleau et al. 2000) ou de gènes spécifiques par bisulfite de sodium et séquençage (Oswald, Engemann et al. 2000).

Les cellules germinales présentent un taux de méthylation d'ADN différent. En effet, le génome d'ovocytes ovulés apparaît moins méthylé que celui de spermatozoïdes.

Dans le zygote, très rapidement après la fertilisation et avant que la réplication d'ADN ne commence, le génome paternel subit une vague de déméthylation, supposée active et impliquant des enzymes de type déméthylases d'ADN ou des protéines à activité ADN glycosylase encore non identifiées. Or l'étude très récente effectuée par Wossidlo et ses collaborateurs (Wossidlo, Arand et al.) a révélé par immunofluorescence en suivant la phosphorylation de la sérine 139 de l'histone H2A.X (γ H2A.X) que ce phénomène était associé à des cassures de l'ADN simple brin. Pendant la réplication de l'ADN, PARP-1, une protéine impliquée dans un mécanisme de réparation de l'ADN simple brin, semble être recrutée par γ H2A.X. Après réplication, une nouvelle vague de déméthylation a lieu même si elle semble moins marquée et l'association PARP-1/ γ H2A.X persiste.

En revanche, le génome maternel ne semble pas impacté par ces deux vagues de déméthylation, une méthylation *de novo* est même observée avant la déméthylation graduelle au cours des divisions cellulaires (donc dite passive). Cependant, des gènes soumis à

l’empreinte parentale, des éléments répétés IAP (Intracisternal A particle) (Lane, Dean et al. 2003) et l’hétérochromatine centromérique (Santos, Hendrich et al. 2002) semblent échapper à toute déméthylation. Après la fécondation, les génomes paternel et maternel retrouvent rapidement un taux de méthylation similaire, ce qui suppose une méthylation rapide *de novo* pour le génome maternel. Puis, une décondensation du pronucleus paternel a lieu. Une faible méthylation d’ADN est associée à une hyperacétylation des histones H4 donc à des changements chromatiniens (Mayer, Niveleau et al. 2000; Oswald, Engemann et al. 2000).

Au stade blastocyste, il y a donc une déméthylation du génome qui en association avec des changements chromatiniens permet l’activation transcriptionnelle de gènes pour aboutir à une spécialisation cellulaire donnant l’embryon et les annexes embryonnaires à partir respectivement du TE (trophectoderme) et ICM (inner cell mass). En effet, des gènes impliqués dans la signalisation et le développement sont sujets ou non au niveau de leur promoteur à des triméthylations des lysines 4 et 27 de l’histone H3 selon le type cellulaire ICM ou TE. Cette association de méthylation sur les 2 sites conduit dans les cellules non différenciées à l’activation transcriptionnelle de ces gènes (Dahl, Reiner et al.). Dans ces mêmes cellules, l’inactivation du chromosome X a déjà commencé (Dvash, Lavon et al.)

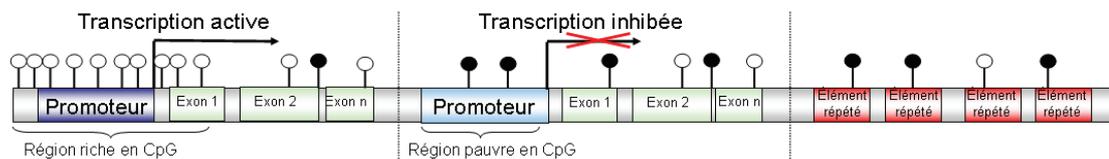
Après ce stade, l’embryon subit une nouvelle vague de méthylation *de novo* pour mettre en place sa propre carte de méthylation alors que le placenta apparaît sous méthylé (Santos, Hendrich et al. 2002). Quant aux cellules germinales primordiales qui donneront par la suite les gamètes, elles paraissent subir une nouvelle vague de déméthylation permettant ainsi une reprogrammation complète pendant la gamétogenèse (Surani 1998) qui se fera différemment selon si la gamète est mâle ou femelle.

IV.B Altérations des profils de méthylation dans les cancers

L’altération des profils de méthylation de l’ADN peuvent conduire à des pathologies comme le syndrome ICF (*Immunodeficiency, Centromeric region instability, Facial anomalies* c’est-à-dire immunodéficiences, instabilité de l’hétérochromatine paracentromérique et dysmorphie faciale) ou participer à la cancérogenèse.

Les cancers sont tous caractérisés par une hypométhylation. Si cette hypométhylation peut toucher des régions géniques spécifiques (De Smet, Lurquin et al. 1999; Pulukuri, Estes et al. 2007), elle concerne bien plus généralement les éléments répétés et on parle alors

Cellule normale



Hyperméthylation locale

Hypométhylation globale
(10-30%)

Cellule cancéreuse

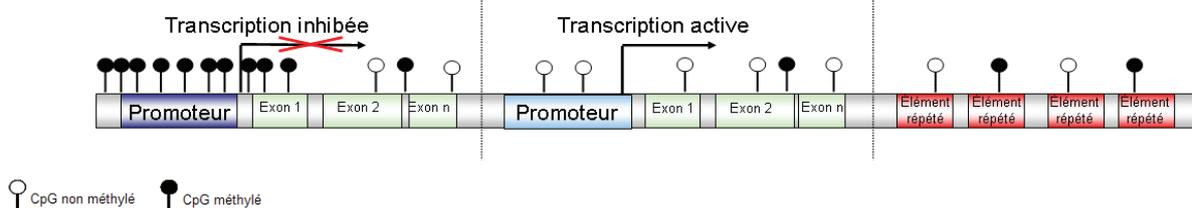


Figure 6 : Altérations des profils de méthylation au cours de la cancérogenèse

Les profils de méthylation de l'ADN de différentes régions (zones riches en CpG, pauvres en CpG, éléments répétés) ont été comparés entre une cellule normale et une cellule cancéreuse. Dans une cellule normale, il apparaît que les promoteurs riches en CpG sont non méthylés et associés à une activation transcriptionnelle alors que les promoteurs pauvres en CpG sont plutôt méthylés et associés à une répression transcriptionnelle. Les éléments répétés sont également les cibles d'une méthylation de l'ADN.

Dans une cellule cancéreuse, une hyperméthylation locale de promoteurs riches en CpG est associée à une hypométhylation globale qui concerne 10 à 30% du génome. Ces deux phénomènes provoquent respectivement l'extinction ectopique de certains gènes notamment des gènes suppresseurs de tumeurs alors que d'autres gènes sont réactivés comme des oncogènes. Les éléments répétés hypométhylés favorisent l'instabilité chromosomique.

d'hypométhylation globale touchant 10 à 30% du génome. Cette dernière est associée à une hyperméthylation locale de régions riches en CpG présentes notamment au niveau de promoteurs de gènes suppresseurs de tumeurs (Ehrlich, Woods et al. 2006) (**Figure 6**).

IV.B.1 Hypométhylation globale du génome

L'hypométhylation globale du génome peut être détectée à des stades assez précoces de la progression tumorale (Kim, Giuliano et al. 1994; Cheng, Schmutte et al. 1997) mais paraît bien plus prononcée au cours de ce processus. Ceci est en corrélation avec les données obtenues suite à la comparaison du taux global de méthylation de nombreuses tumeurs métastatiques humaines (mélanome, tumeurs du sein, pancréas, col de l'utérus, peau, ...), tumeurs bénignes (du col de l'utérus, sein, ovaires) et tissus normaux évalué par chromatographie en phase liquide à haute performance (Feinberg, Gehrke et al. 1988). Seules les tumeurs malignes semblent présenter une hypométhylation. Elle ne correspond pas à une diminution du taux de cytosines qui aurait pu être expliquée par un défaut dans la réparation de l'ADN des désaminations spontanées de cytosines méthylées en thymine (Gama-Sosa, Slagel et al. 1983). Ces mêmes observations ont été obtenues en comparant de manière plus stricte les tissus métastatiques à leurs contre-parties normales ou bénignes pour le cancer du sein par exemple (Soares, Pinto et al. 1999) ou suite à l'utilisation de cellules souches mésenchymales primaires immortalisées puis progressivement transformées. L'étude de la méthylation de différentes séquences répétées reflétant la méthylation globale du génome, par une approche globale (MethLight) et par bisulfite suivi d'un séquençage, a montré qu'une hypométhylation apparaît seulement au dernier stade de transformation et qu'elle prend graduellement de l'importance (Wild, Funes et al.). Ils ont cependant établi que cette hypométhylation ne constitue pas un pré-requis pour induire la transformation cellulaire.

L'hypométhylation semble localisée dans des zones assez dispersées dans les tissus cancéreux (Shann, Cheng et al. 2008). Elle marque essentiellement les séquences répétées type *Satellite-2 (Sat-2)* (Narayan, Ji et al. 1998) ou *Alu* (Rodriguez, Vives et al. 2008), et est associée à une réexpression ectopique de certains gènes comme les proto-oncogènes (Badal, Chuang et al. 2003), les rétrotransposons ou les gènes soumis à l'empreinte ainsi qu'à une instabilité chromosomique (Lengauer, Kinzler et al. 1997; Schulz, Elo et al. 2002). Elle semble également intervenir dans l'élongation transcriptionnelle.

L'hypométhylation globale du génome apparaît être un phénomène indépendant de l'hyperméthylation locale de promoteurs de gènes spécifiques (Ehrlich, Woods et al. 2006).

IV.B.2 Hyperméthylation locale des promoteurs riches en CpG

L'hyperméthylation locale est retrouvée au niveau de promoteurs riches en CpG (Smiraglia, Rush et al. 2001). Cette hyperméthylation correspond à un phénomène précoce dans de nombreux cancers (Kanai and Hirohashi 2007). Si au départ l'hyperméthylation se traduit par une mosaïque de CpG méthylés et non méthylés, elle couvre plus de CpG au fur et à mesure de la progression tumorale. Au final, ces régions seront complètement méthylées ou encore clairsemées de CpG non méthylés (Hodges, Smith et al. 2009).

Dans les cancers hépatiques, le promoteur du gène codant l'E-cadhérine impliquée dans les jonctions cellulaires ou *HIC-1* (*hypermethylated-in-cancer 1*), un gène suppresseur de tumeurs (Chen, Zeng et al. 2003), subissent une hyperméthylation progressive corrélée à une perte d'expression de plus en plus importante ainsi que pour l'E-cadhérine à une perte d'adhésion cellulaire, une marque caractéristique des cellules cancéreuses (Yoshiura, Kanai et al. 1995). Cette hyperméthylation peut également marquer des points chauds pour des pertes de régions chromosomiques, participant ainsi une instabilité chromosomique.

De nombreux gènes suppresseurs de tumeurs ont été identifiés comme étant hyperméthylés dans des cancers de façon spécifique ou aspécifique aux tissus. A titre d'exemples, le promoteur du gène *p16* est retrouvé hyperméthylé dans de nombreux cancers (Ruas and Peters 1998) et celui de *p14* hyperméthylé dans les cancers du colon (Esteller, Tortola et al. 2000). Chaque type de cancer paraît posséder son propre groupe de gènes suppresseurs de tumeurs hyperméthylés, ce qui a donné naissance à l'appellation « phénotype méthylome » CIMP (CpG island methylator phenotype) (Issa 2004). En parallèle à ces groupes, certains gènes codant pour les microARN paraissent réprimés par ce mécanisme (Liu, Teng et al.), un « phénotype miRNome » pourrait donc voir le jour (Davalos and Esteller). En revanche, le promoteur *TIMP-3* est hyperméthylé dans certains cancers rénaux (Masson, Rioux-Leclercq et al.) et colorectaux (Rhee, Bachman et al. 2002) et hypométhylé dans des leucémies myéloïdes aiguës (Rohrs, Dirks et al. 2009). L'hyperméthylation des promoteurs de gènes suppresseurs de tumeur n'est donc pas un phénomène systématique.

Les gènes dont les promoteurs deviennent hyperméthylés dans les cancers codent pour des microARN ou protéines impliqués dans de nombreux mécanismes. En effet, p14, p15 et p16 sont impliquées dans le contrôle du cycle cellulaire, MLH1 et MGMT (O6-méthylguanine methyltransferase) interviennent dans la réparation de l'ADN (Weller, Stupp et al.; Fleisher, Esteller et al. 2001), l'E-cadhérine et CDH13 dans l'adhésion cellulaire (Kim, Han et al. 2005). Des facteurs de transcription comme GATA-4 et GATA-5 peuvent aussi voir leurs expressions réprimées (Wen, Akiyama et al.). Enfin, miR-184 régule le gène *Numbl* dont la protéine est impliquée dans la prolifération des cellules souches neurales (Liu, Teng et al.), d'où un rôle indirect des microARN via la régulation de leurs cibles.

IV.B.3 Implication des DNMT dans ces altérations de profils

IV.B.3.a Dérégulations de l'expression des DNMT dans les cancers

IV.B.3.a.i Mise en évidence de ces dérégulations

Etant donné que les tumeurs malignes présentent des profils de méthylation altérés, il a été proposé que les taux des protéines catalysant la méthylation de l'ADN soient dérégulés au cours de la progression tumorale et responsables des profils observés. Une étude réalisée au cours de l'embryogenèse suggérait que la concentration de DNMT ne déterminait pas systématiquement le taux de méthylation de l'ADN (Trasler, Trasler et al. 1996). Cette hypothèse est renforcée par les travaux réalisés par Vertino et ses collaborateurs. Ils ont comparé, par digestion enzymatique sensible à la méthylation et révélée par Southern Blot, la méthylation d'îlots CpG endogènes présents dans des cellules fibroblastiques embryonnaires pulmonaires (IMR-90) immortalisées par l'introduction de SV40 et transfectées par un plasmide contenant l'ADNc de *DNMT1* ou un plasmide vide. Une hyperméthylation de plus en plus importante a lieu en fonction du temps et n'est pas corrélée au taux d'expression de DNMT1 (Vertino, Yen et al. 1996).

Par contre, par la suite, il a été montré dans les tissus cancéreux que les taux d'ARNm *DNMT1*, *DNMT3a* et *DNMT3b* paraissaient plus importants que ceux quantifiés dans les tissus normaux (Robertson, Uzvolgyi et al. 1999; Mizuno, Chijiwa et al. 2001). Ces observations sont corrélées à d'autres données obtenues suite à l'étude d'ARNm de *DNMT1* même si l'écart entre les tissus normaux versus cancéreux apparaît très variable. Ceci

s'expliquerait par des variations intrinsèques aux tissus cancéreux et par des fluctuations, visualisées au niveau protéique par immunohistochimie, entre les pourcentages de cellules exprimant faiblement, plus fortement et très fortement *DNMT1* au cours de progression tumorale (De Marzo, Marchi et al. 1999) avec une résultante globale d'augmentation du taux protéique de DNMT1 (Saito, Kanai et al. 2003; Peng, Kanai et al. 2005; Sawada, Kanai et al. 2007). Ces arguments pourraient être avancés pour DNMT3a ou DNMT3b. L'épissage alternatif pourrait également expliquer ces observations selon si la forme prédominante a ou non une activité catalytique (Eads, Danenberg et al. 1999; Saito, Kanai et al. 2002).

IV.B.3.a.ii Conséquences fonctionnelles de ces dérégulations

Les fluctuations des taux protéiques de DNMT1, DNMT3a et DNMT3b au cours de la progression tumorale pourraient être directement responsables des profils de méthylation observés. En effet, la comparaison par approche globale des patrons de méthylation entre les cellules ES murines et les cellules ES invalidées pour *Dnmt1* et/ou *Dnmt3a* et/ou *Dnmt3b* (Hattori, Abe et al. 2004) a suggéré que *Dnmt1* est surtout impliquée dans la méthylation des îlots de CpG alors que DNMT3a et b interviennent plutôt au niveau des séquences répétées. En accord avec ces données, des interférences à l'ARN dirigés contre *DNMT1* dans des cellules humaines cancéreuses pulmonaires et mammaires ont montré pour des gènes suppresseurs de tumeurs une diminution de la méthylation au niveau de leurs promoteurs contenant des îlots de CpG initialement hyperméthylés dans ces lignées ainsi qu'une expression de ces gènes (Suzuki, Sunaga et al. 2004). Une étude similaire menée dans les lignées ovariennes vient conforter ces résultats. L'invalidation de DNMT3b semble avoir moins d'impact sur cette hyperméthylation (Leu, Rahmatpanah et al. 2003).

Cependant, des résultats inverses c'est-à-dire une conservation de l'hyperméthylation des îlots de CpG est retrouvée dans une lignée colorectale suite à l'invalidation de DNMT1 transitoire ou stable par interférence à l'ARN (Ting, Jair et al. 2004) ou par recombinaison homologue (Rhee, Bachman et al. 2002). En effet, dans cette dernière étude, l'invalidation de *DNMT3B* ou *DNMT1* seule n'a aucun impact sur la méthylation de gènes hyperméthylés comme *p16* ou *TIMP-3* alors que leur combinaison permet l'hypométhylation d'îlots de CpG, de séquences répétées et de gènes soumis à l'empreinte parentale. Ces données laissent donc penser que les techniques utilisées pour l'invalidation des DNMT ont une influence directe

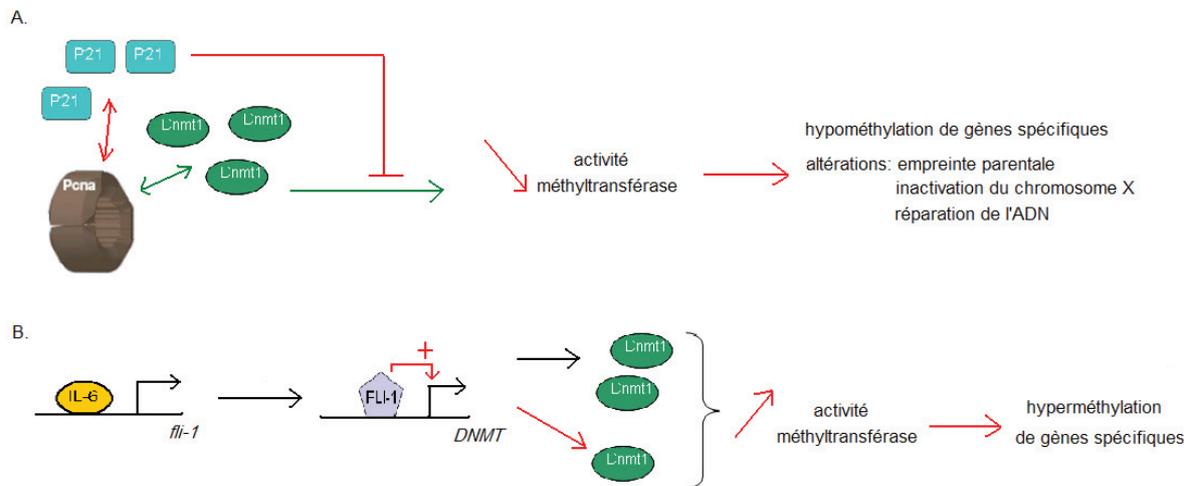


Figure 7 : Influence de l'environnement sur l'activité méthyltransférase d'ADN

- A) Le taux de P21 peut augmenter au cours de la cancérogenèse. A ce moment-là, cette protéine peut entrer en compétition avec DNMT1 pour la fixation à PCNA. Plus la compétition est en faveur de P21 moins DNMT1 peut exercer son activité méthyltransférase d'ADN. Ceci donne lieu à une hypométhylation qui peut toucher des gènes spécifiques et/ou altérer l'empreinte parentale, l'inactivation du chromosome X et perturber la réparation de l'ADN.
- B) La cytokine IL-6 peut être sécrétée par des cellules érythrocytaires, fixée le promoteur de *fli-1*. La protéine alors produite favoriserait l'expression de DNMT1 et par conséquent augmenterait l'activité méthyltransférase d'ADN, engendrant une hyperméthylation locale.

sur les résultats observés ou plus probablement que ces observations reflètent une spécificité de l'environnement cellulaire.

L'hypométhylation serait, quant à elle, expliquée au moins en partie par l'altération de la transcription de *DNMT3b*. En effet, jusqu'à l'étude réalisée par Saito et ses collaborateurs, les données concernant l'expression de DNMT3b augmentée dans les cancers (Robertson, Uzvolgyi et al. 1999; Saito, Kanai et al. 2002) contredisaient les résultats obtenus suite à l'inactivation de *Dnmt3b* conduisant une instabilité chromosomique due à l'hypométhylation de régions satellites péri-centromériques (Xu, Bestor et al. 1999). En fait, l'épissage alternatif de *DNMT3b* serait augmenté vu que près de 20 variants ont été identifiés dans les cancers et serait en faveur de formes telles que DNMT3b4 ne possédant pas de site catalytique (Saito, Kanai et al. 2002; Ostler, Davis et al. 2007). Ces variants inactifs joueraient en plus un rôle négatif dans la régulation de l'activité méthyltransférase.

IV.B.3.b Influence de l'environnement sur l'activité des DNMT

L'environnement paraît influencer une reprogrammation des profils de méthylation altérés (Ron-Bigger, Bar-Nur et al.; Costa, Seftor et al. 2009). Il pourrait passer par une perturbation de l'activité des DNMT.

Ainsi, dans les cellules humaines colorectales normales, il y existe un fort différentiel de concentration entre les protéines p21 et DNMT1, cette dernière étant bien plus exprimée, alors que dans ces cellules devenues cancéreuses, une co-expression de DNMT1 et p21 a été observée (De Marzo, Marchi et al. 1999). Or, PCNA peut aussi interagir avec p21. Une compétition entre p21 et DNMT1 pour la liaison à PCNA dans les foyers de réplication aurait alors lieu et perturberait la maintenance de la méthylation de l'ADN par DNMT1 (Chuang, Ian et al. 1997). Cela donnerait une hypométhylation de gènes spécifiques et/ou aurait des répercussions sur les gènes soumis à l'empreinte, l'inactivation du chromosome X, et la réparation de l'ADN ce qui participerait à l'instabilité chromosomique (**Figure 7A**).

A l'inverse, DNMT1 pourrait voir son activité de maintenance de l'ADN favorisée de façon directe ou indirecte. En effet, DNMT1 paraît interagir avec les protéines E1A et E7, provenant d'adénovirus responsables de cancers du col de l'utérus, ce qui stimulerait son activité méthyltransférase d'ADN *in vitro* (Burgers, Blanchon et al. 2007). De plus, la cytokine pro-inflammatoire IL-6, sécrétée par des cellules érythrocytaires, semblerait fixer le

promoteur de *fli-1* qui serait activé et la protéine FLI-1 viendrait à son tour se lier sur le promoteur de *DNMT1*, augmentant l'expression de DNMT1 et son activité (Hodge, Xiao et al. 2001) (**Figure 7B**). Des cellules donnant des leucémies myéloïdes multiples, sensibilisées à un environnement contenant IL-6, voient le promoteur de *p53* méthylé et le gène non exprimé, a priori à cause de l'augmentation de l'activité de DNMT1. L'ajout de zébularine, un inhibiteur des DNMT, permettrait d'abolir cette méthylation donc *p53* serait réexprimée et l'apoptose augmentée (Hodge, Peng et al. 2005).

IV.C Conclusion

La méthylation de l'ADN correspond à une marque mise en place et maintenue par les DNMT1, 3a et 3b. Elle intervient dans de nombreux processus biologiques indispensables chez les vertébrés. Cette marque apparaît réversible, avec des phénomènes dissociés d'hypo- et d'hyperméthylation. Ces altérations de profils de méthylation de l'ADN, dont les mécanismes restent majoritairement à élucider, contribuent notamment à la cancérogenèse.

La méthylation de l'ADN a pour rôle majeur d'initier la répression transcriptionnelle passant par des changements chromatinien. Aussi, au cours du chapitre suivant, nous allons voir comment les protéines liant l'ADN méthylé jouent les intermédiaires entre ces mécanismes en « lisant et traduisant » ce signal.

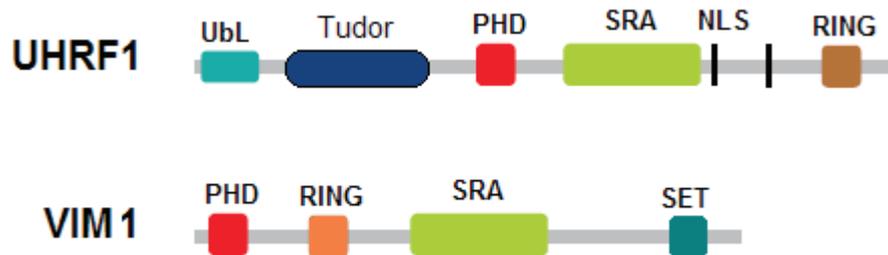


Figure 8: Schématisation protéique des membres fondateurs des deux familles de protéines à domaine SRA

Il existe deux familles de protéines à domaine SRA.

La première est la famille « UHRF » (Ubiquitin like containing plant homolog domain (PHD) and RING finger domains) dont le membre fondateur est UHRF1. Cette protéine comprend les domaines UbL (Ubiquitine-like), Tudor, PHD (Plant Homolog Domain), SRA (SET and Ring finger associated), 2 domaines NLS (Nuclear localization signal) et RING.

La seconde est la famille VIM (Variant in methylation) présente uniquement chez les plantes. Le membre fondateur VIM1 se distingue notamment de la protéine UHRF1 car il contient un domaine SET en plus des domaines PHD, RING et SRA.

Une lecture par les protéines liant l'ADN méthylé

Il existe trois familles de protéines ayant la propriété de lier des CpG méthylés et qui se distinguent par leurs domaines de reconnaissance à l'ADN méthylé. La première famille identifiée a été celle des protéines à methyl-CpG-binding domain (MBD). La famille Kaiso appartenant aux protéines à doigts de zinc (ZF) puis la famille à domaine SRA (SET and Ring finger associated) principalement étudiée chez les plantes, ont ensuite été caractérisées.

I- La famille à domaine SRA

Les protéines à domaine SRA (SET and Ring finger associated aussi appelé YGD) se divisent en deux groupes.

Le premier groupe est composé de deux membres chez l'Homme : les protéines nucléaires UHRF1 (Ubiquitin like containing plant homolog domain (PHD) and RING finger domains 1) aussi appelée ICBP90 (inverted CCAAT box binding protein of 90 kDa) ou NP95, et UHRF2 également nommée NIRF (Np95/ICBP90-like ring finger) (**Figure 8**). Leurs orthologues murins correspondent respectivement à Np95 et Np97. La protéine UHRF1 contient un domaine Ubiquitin like (Ubl) qui lui confère, d'après les études menées sur Np95, une activité d'ubiquitylation surtout des histones H3 (Citterio, Papait et al. 2004), un domaine leucine zipper, deux domaines en doigts de zinc, un de type PHD reconnaissant les histones méthylées sur leur lysine 4 et l'autre de type RING dirigeant la dégradation des protéines cibles par le protéasome, deux domaines de signal de localisation nucléaire (NLS) et un domaine SRA qui lui permet de lier l'ADN méthylé sans spécificité de séquence (Unoki, Nishidate et al. 2004) ou avec plus d'affinité l'ADN hémi-méthylé (Bostick, Kim et al. 2007; Sharif, Muto et al. 2007). Récemment un domaine en tandem Tudor a été identifié entre Ubl et PHD et paraît former une poche pour interagir avec la queue de l'histone H3 triméthylée en lysine 9 (Rottach, Frauer et al.). Cette protéine semble impliquée dans la prolifération cellulaire (Unoki, Nishidate et al. 2004) avec un rôle essentiel dans la transition G1/S (Arima, Hirota et al. 2004). De plus, cette protéine semble intervenir dans la maintenance de la méthylation de l'ADN en guidant DNMT1, recrutée par PCNA au niveau des fourches de

réplication, vers les sites hémi-méthylés (Bostick, Kim et al. 2007; Sharif, Muto et al. 2007; Achour, Jacq et al. 2008) à travers un complexe formé de NP95, de DNMT1, de HDAC1 (Unoki, Nishidate et al. 2004) et de G9a (Esteve, Chin et al. 2006) qui interviendrait dans les di- ou triméthylations de la lysine 9 de l'histone H3. De plus, comme le montre des ChIP, cette protéine semble participer à la répression transcriptionnelle de gènes suppresseurs de tumeurs comme *p14*, *p16*, *RAR β* (Unoki, Nishidate et al. 2004), *RBI* (Jeanblanc, Mousli et al. 2005), la *cycline E* (Nielsen, Schneider et al. 2001) et *p21* (Kim, Esteve et al. 2009) pour lequel ce complexe est recruté au niveau du promoteur par Gfi1. Les multiples interactions d'ICBP90 associées à son intervention dans l'ubiquitylation des histones et dans la formation de l'hétérochromatine (Karagianni, Amazit et al. 2008), font de cette protéine un lien direct entre la méthylation de l'ADN, les modifications des histones et de la structure chromatinienne. Dans les cellules tumorales, elle semble surexprimée (Unoki, Nishidate et al. 2004; Achour, Jacq et al. 2008) et son taux ne plus être dépendant du cycle cellulaire. Elle paraît alors également jouer un rôle dans l'angiogenèse en stimulant indirectement, en collaboration avec DNMT1, l'expression de *VEGF* via son inhibition renforcée de *p16* (Achour, Jacq et al. 2008). Moins de données sont disponibles pour la protéine UHRF2. Néanmoins, il apparaît qu'elle contient les mêmes domaines, est aussi impliquée la régulation du cycle cellulaire, participe à la transition G1/S et a une activité d'ubiquitylation.

Les plantes *Arabidopsis thaliana* sont les seules à compter en plus de ce groupe comportant pour elles cinq membres appelés VIM (Variant in methylation) jouant un rôle partiellement redondant au moins pour VIM1 (**Figure 8**), 2 et 3 (Woo, Dittmer et al. 2008), un second groupe incluant essentiellement les histones méthyltransférases d'histones SUVH. Ces protéines contiennent en plus du domaine SRA un domaine SET, impliqué dans la méthylation des histones. Récemment, il a été découvert que SUVH2 et 9 interagissaient avec DRM2, l'orthologue des DNMT3, leur conférant un rôle dans la méthylation de l'ADN (Johnson, Law et al. 2008). Les protéines à domaine SRA semblent ainsi avoir un rôle majeur dans la méthylation de l'ADN *de novo* et de maintenance.

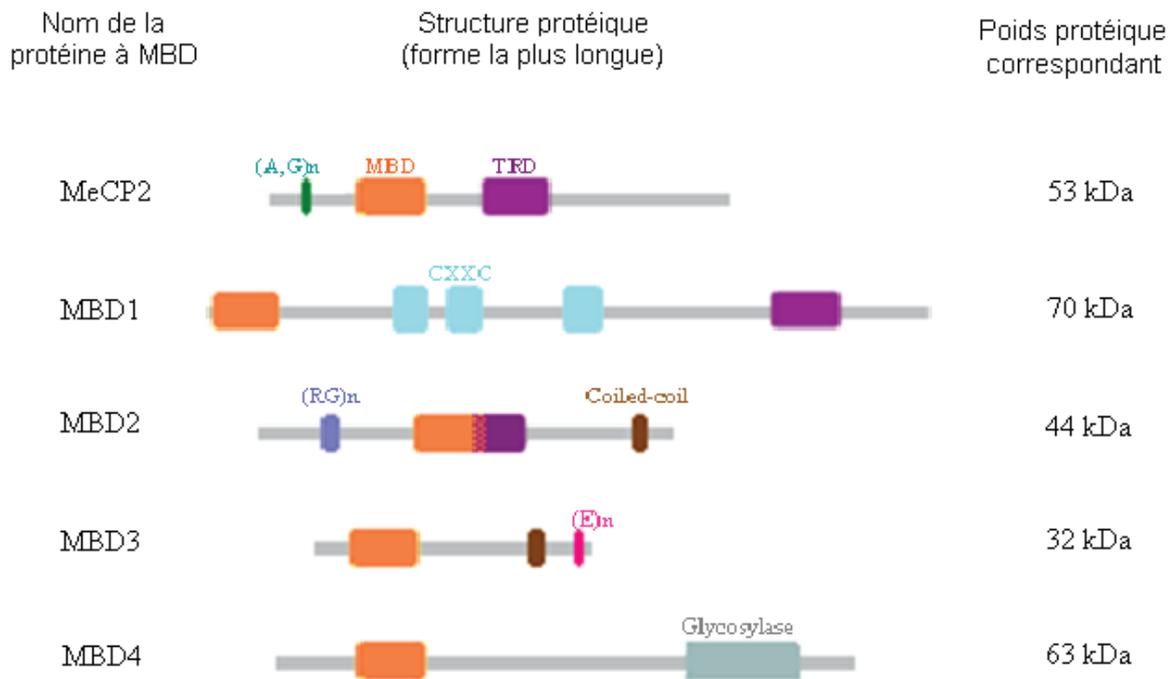


Figure 9 : Schématisation protéique de la famille des protéines à « methyl-CpG binding domain » (protéines à MBD)

La famille de protéines à MBD comprend 5 membres : MeCP2, le membre fondateur, MBD1, MBD2, MBD3 et MBD4. Ce schéma représente les formes les plus longues et leur poids protéique associé (en kilodaltons). Ces protéines comprennent toutes un domaine MBD. MeCP2, MBD1 et MBD2 partagent également un domaine de répression transcriptionnelle (TRD). Elles comprennent aussi des domaines spécifiques : CXXC pour MBD1, glycosylase pour MBD4, coiled-coil pour MBD2 et MBD3. MeCP2, MBD2 et MBD3 ont des zones présentant des motifs d'acides aminés répétés avec A, Alanine ; G, glycine ; E, acide glutamique ; n, répétition du motif.

II- La famille des protéines à methyl-CpG-binding domain (MBD)

II.A Caractérisation des domaines peptidiques des protéines à MBD

II.A.1 Un domaine de liaison à l'ADN méthylé (ou MBD)

L'analyse des domaines fonctionnels de la protéine MeCP2, membre fondateur de la famille, a permis d'identifier une région particulière responsable de la liaison avec les CpG méthylés, le « Methyl-CpG Binding Domain » ou MBD (Meehan, Lewis et al. 1992) (**Figure 9**). Puis, à partir de recherches d'homologies de séquences pour cette région dans des banques d'EST (expressed sequence tags) humaines les autres protéines à MBD -MBD1, MBD2, MBD3 et MBD4- ont été découvertes (Hendrich and Bird 1998).

Ce domaine MBD constitue la seule homologie de séquences entre les différents membres de la famille, bien que les protéines MBD2 et MBD3 présentent plus de 70% de similitudes. Il correspond à une séquence d'environ 70 acides aminés formant une structure en sandwich α/β . Il permet, *in vitro* comme *in vivo*, la liaison à l'ADN hémi-méthylé ou à un seul CpG méthylé (Hendrich and Bird 1998), hormis pour la protéine MBD3 chez les mammifères. Cette exception s'explique par les substitutions identifiées dans le MBD de MBD3 (Fraga, Ballestar et al. 2003) qui le rendent non fonctionnel. Néanmoins, MBD3 peut être retrouvée au niveau de l'ADN méthylé grâce à son association avec MBD2 (Tatematsu, Yamazaki et al. 2000).

II.A.2 Le domaine de répression transcriptionnelle (TRD)

Des expériences de délétions du gène *MeCP2* fusionné avec le domaine de liaison à l'ADN de *Gal4* (Nan, Campoy et al. 1997) ont permis d'identifier un domaine de MeCP2 impliqué dans la répression transcriptionnelle. Il s'agit du domaine de répression transcriptionnelle ou TRD que possèdent toutes les protéines à MBD, à l'exception de MBD3. Si les domaines TRD des protéines à MBD n'ont aucune similarité nucléotidiques ou peptidiques, ils interviennent de la même manière en recrutant des complexes enzymatiques capables de modifier la structure de la chromatine.

II.A.3 Les autres domaines des protéines à MBD

Les protéines à MBD sont également constituées d'autres domaines leur conférant des activités spécifiques.

Ainsi, la protéine MBD1 a la propriété de se lier à l'ADN non méthylé. En effet, elle contient, selon les transcrits (Fujita, Takebayashi et al. 1999), 2 à 3 domaines CXXC c'est-à-dire des motifs en doigt de zinc. Seul le domaine CXXC3 présent dans 3 transcrits (variants 1, 2 et PCM1) assure une liaison nucléotidique en absence de méthylation de l'ADN (Jorgensen, Ben-Porath et al. 2004).

Quant à la protéine MBD4 ou MED1, elle semble reconnaître des mésappariements de l'ADN dus à la méthylation par erreur en O6 de guanines (8-oxoG) et posséder un domaine glycosylase lui permettant d'intervenir dans la réparation de l'ADN (Bellacosa, Cicchillitti et al. 1999).

II.B Localisations nucléaire et tissulaire des protéines à MBD

II.B.1 Une localisation nucléaire qui coïncide avec la présence d'ADN méthylé

Les protéines à MBD sont des protéines exclusivement nucléaires présentes au niveau des 5-methylcytosines. En effet, des transfections transitoires de plasmides codant ces protéines et contenant des séquences cibles pour une visualisation par immunofluorescence ont été réalisées dans des cellules murines. Ces protéines sont alors très majoritairement retrouvées au niveau des foyers d'hétérochromatine des régions péricentromériques riches en CpG méthylés, à l'exception de MBD3 (Nan, Tate et al. 1996; Hendrich and Bird 1998) qui ne reconnaît pas l'ADN méthylé. Une petite partie de ces protéines peut quelquefois être localisée au niveau d'autres régions méthylées avec une densité en CpG moins importante, à savoir l'euchromatine et les séquences satellites (Fujita, Takebayashi et al. 1999).

II.B.2 Une expression tissulaire quasiment ubiquitaire

Dans les cellules souches embryonnaires, la protéine MBD3 est fortement exprimée alors que les autres protéines à MBD sont très peu détectées (Hendrich and Bird 1998). Ces données concordent avec les invalidations des gènes codant pour ces protéines puisqu'un phénotype léthal est observé uniquement pour les souris *Mbd3*^{-/-} suite à la dysfonction d'un complexe répresseur auquel appartient cette protéine (Guy, Hendrich et al. 2001; Hendrich, Guy et al. 2001; Zhao, Ueba et al. 2003). En revanche, si *Mbd2* et *Mbd4* sont détectées dès le stade blastula, *Mecp2* l'est plus tardivement (Kantor, Makedonski et al. 2003; Huntriss, Hinkins et al. 2004).

Dans les cellules somatiques, toutes les protéines à MBD sont ubiquitaires avec cependant une différence de distribution selon les tissus (Hendrich and Bird 1998). En effet, la protéine MeCP2 est peu exprimée dans les testicules de rat alors qu'elle est relativement abondante dans le cerveau chez l'Homme, la souris et le rat. Etant donné son rôle important dans le syndrome de Rett (Amir, Van den Veyver et al. 1999), une maladie neurodégénérative touchant les filles, des études plus complètes des différentes régions du cerveau ont été réalisées. Il apparaît que le bulbe olfactif, le cervelet, le cortex préfrontal dans le striatum, le septum et l'hippocampe présentent des taux importants de *Mecp2* chez le rat (Cassel, Revel et al. 2004), la souris (Urduingio, Lopez-Serra et al. 2008). Plus généralement, MeCP2 s'exprime plus fortement dans les neurones matures post-mitotiques qui sont des cellules différenciées (Skene, Illingworth et al.; Balmer, Goldstine et al. 2003; Kishi and Macklis 2004), que dans les cellules se divisant activement ou en culture. D'autre part, on peut distinguer des niveaux d'expression différents selon les transcrits. Ainsi, MeCP2 alpha prédomine dans le cerveau, le thymus et les poumons de souris (Dragich, Kim et al. 2007).

De plus, si MBD1 et MBD3 présentent un profil d'expression assez similaire à MeCP2 dans les tissus, MBD1 est plus fortement exprimée dans les cellules en culture. Quant à MBD2, elle aurait un patron d'expression assez opposé à MeCP2 avec par exemple une faible expression dans le cerveau et un taux d'ARN messager 20 fois plus important que celui de *MeCP2* dans les cellules HeLa (Auriol, Billard et al. 2005) ou tissus mammaires.

Ainsi, les concentrations des protéines à MBD diffèrent selon le type cellulaire et le stade de développement.

II.C Une famille impliquée dans la répression transcriptionnelle

Il a été démontré que tous les membres de la famille des protéines à MBD participent à la répression transcriptionnelle, même si comme nous le verrons au cours de ce chapitre leurs rôles et leurs cibles sont différents.

II.C.1 Des répresseurs transcriptionnels dépendants de la méthylation

La première mise en évidence d'un rôle transcriptionnel des protéines à domaine liant les CpG méthylés a été faite avec l'étude du complexe MeCP1 formé par MBD2 et Mi2/NuRD contenant MBD3. Puis, des expériences de transfections, dans des cellules déficientes ou non en MBD2, de plasmides méthylés ou non et contenant le gène luciférase (Hendrich, Guy et al. 2001) ont permis de confirmer le rôle répresseur de MBD2. De plus, lors d'expériences de transcription *in vitro* la fonction répressive méthylation-dépendante de la protéine MeCP2 a aussi été mise en évidence (Meehan, Lewis et al. 1992). Des fusions de fragments de *Mbd1* avec l'ADN de *Gal4* couplés au promoteur de la β -actine révèlent également un rôle de répression pour MBD1 (Ng, Jeppesen et al. 2000). Enfin, une seule étude (Kondo, Gu et al. 2005) utilisant ce même type d'approche suggère que MBD4 agit comme répresseur.

Ainsi, les protéines MeCP2, MBD1, MBD2 et sans doute MBD4 forment une famille de répresseurs transcriptionnels dépendants de la méthylation. Le rôle de MBD3 dans l'inhibition transcriptionnelle, quel que soit l'état de méthylation de l'ADN, ne semble dû qu'à sa présence dans le complexe Mi2/NuRD, faisant de lui un acteur indirect. Cela permettrait à MBD3 de participer au contrôle de la transition G₂/M au cours du cycle cellulaire (Noh, Lim et al. 2009).

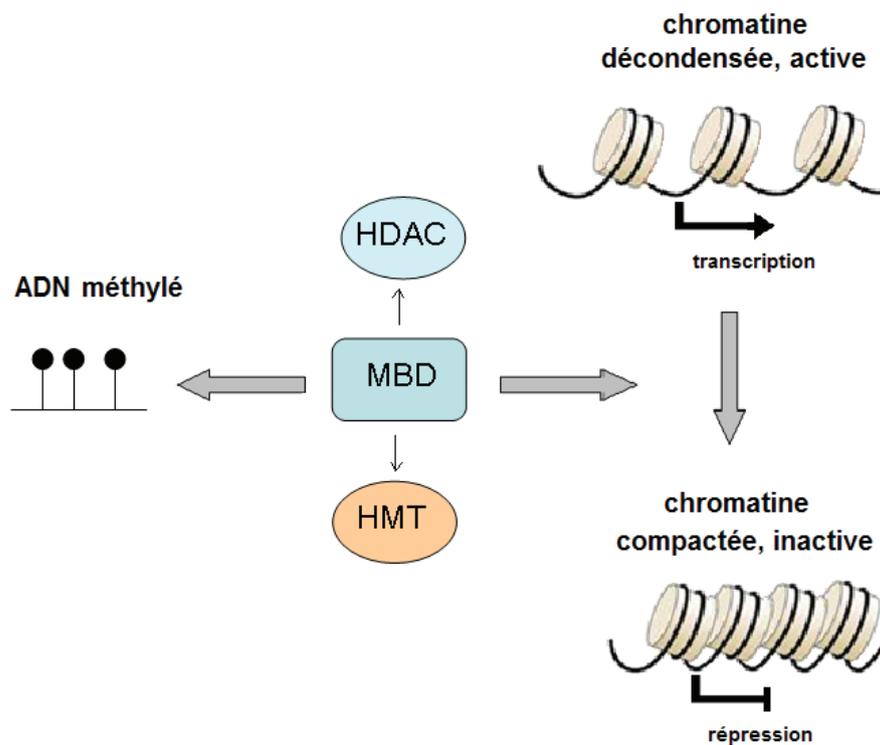


Figure 10 : Les protéines à MBD font le lien entre méthylation de l'ADN, répression transcriptionnelle et compaction chromatinienne

La répression transcriptionnelle dépendante de la méthylation de l'ADN passe au moins en partie par les protéines à MBD qui recrutent des désacétylases d'histones (HDAC) et/ou des méthyltransférases d'histones (HMT) afin de compacter la chromatine.

II.C.1.a Une répression dépendante de la méthylation associée à une compaction chromatinienne

De nombreuses études montrent que la répression transcriptionnelle par les protéines à MBD passe, au moins en partie, par le recrutement de désacétylases d'histones (HDAC) et/ou de méthyltransférases d'histones (HMT) afin de compacter la chromatine (**Figure 10**).

II.C.1.a.i Les protéines à MBD initient une répression dépendante de la méthylation de l'ADN en recrutant des désacétylases d'histones

- **MeCP2, premier répresseur de la famille identifié comme recruteur de HDAC**

L'équipe du Pr. A. Bird (Hendrich and Bird 1998) montraient une co-localisation de MeCP2 et des histones hypoacétylés, ce qui suggérait une éventuelle interaction entre MeCP2 et les complexes enzymatiques contenant des désacétylases d'histones (HDAC). Cette même équipe (Nan, Ng et al. 1998) a fusionné la protéine MeCP2 à la glutathione S-transférase (GST) et ciblée ce complexe par différents anticorps. Ils ont ainsi révélé que MeCP2 via son domaine TRD interagissait avec le co-répresseur mSin3A et les désacétylases d'histones HDAC1 et 2. D'autres études mettent en évidence l'interaction de MeCP2 avec les HDAC et Sin3A (Martinowich, Hattori et al. 2003) auxquels peut s'ajouter le complexe REST/RNSF (RE1-silencing transcription factor/Neural restrictive silencing factor) (Naruse, Aoki et al. 1999).

Les désacétylases d'histones catalysent l'hydrolyse des groupements acétyles des histones H3 et H4 et entraînent la compaction de la chromatine. Ainsi, ces expériences ont permis de mettre en évidence pour la première fois le lien direct existant entre la méthylation de l'ADN et les changements de la chromatine.

D'autres complexes HDAC sont associés à MeCP2. Ainsi, des expériences de chromatographie d'affinité « GST-pull down » montrent *in vitro* que cette protéine à MBD lie d'autres complexes co-répresseurs, c-Ski et N-CoR, via les domaines TRD et MBD respectivement (Kokura, Kaul et al. 2001). Des co-immunoprécipitations de protéines

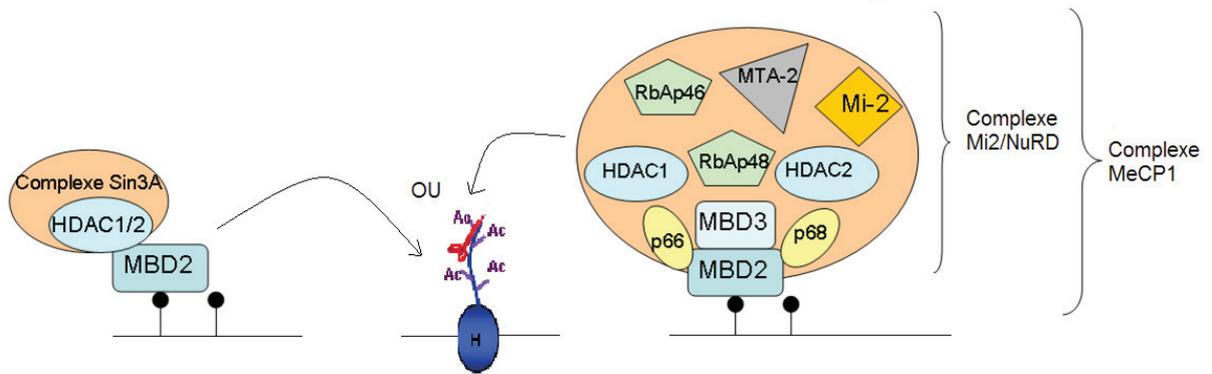


Figure 11 : Le répresseur MBD2 agit en recrutant des complexes contenant des HDAC suite à sa liaison sur l'ADN méthylé

Après liaison sur l'ADN méthylé, la protéine MBD2 peut recruter le complexe co-répresseur Sin3 ou le complexe Mi2/NuRD qui contiennent notamment les désacétylases d'histones HDAC1 et 2 qui vont alors désacétyler les lysines des histones présentes sur cette région.

confirment l'interaction *in vivo* de MeCP2 avec c-Ski qui paraît intervenir dans la répression transcriptionnelle.

- **MBD2 a une fonction répressive très HDAC-dépendante**

MBD2 s'associe très majoritairement à différents complexes répresseurs contenant des désacétylases d'histones. En effet, elle appartient au complexe MeCP1 (Ng, Zhang et al. 1999) et au complexe NuRD (Zhang, Ng et al. 1999) qui ne forment en réalité qu'un seul complexe. MeCP1 regroupe les unités protéiques du complexe Mi2/NuRD, à savoir HDAC1, HDAC2, RbAp46, RbAp48, l'ATPase Mi-2, MTA2 (Metastasis associated 1 family, member 2) et MBD3, associé à MBD2, et en plus p66 et p68, 2 isoformes d'une protéine à doigts de zinc qui interagit avec MBD2 (Feng and Zhang 2001). De plus, des expériences de double hybride et de co-sédimentation mettent en évidence que la protéine MBD2 peut aussi recruter le complexe répresseur Sin3A (Boeke, Ammerpohl et al. 2000) (**Figure 11**).

Des analogues de MBD3, MBD3L1 exprimé uniquement dans les spermatides et MBD3L2 plus ubiquitaire, ne possédant pas de domaine MBD mais un domaine « coiled-coil », paraissent pour voir remplacer MBD3 et MBD2 respectivement dans le complexe Mi-2/NuRD et ainsi participer à la répression transcriptionnelle (Jiang, Jin et al. 2004; Jin, Jiang et al. 2005). Une seule étude paraît démontrer que les complexes Mi-2/NuRD contiennent de manière exclusive MBD2 ou MBD3 ainsi que d'autres protéines (Le Guezennec, Vermeulen et al. 2006) comme DOC-1 (Spruijt, Bartels et al.) et plus spécifiquement la méthyltransférase d'arginine PRMT5 pour le complexe MBD2/Mi-2/NuRD. Ces données n'appuient pas les conclusions faites suite aux invalidations des gènes *Mbd2* et *Mbd3* (Hendrich, Guy et al. 2001). En effet, les doubles mutants *Mbd2*^{-/-} et *Mbd3*^{+/-} présentent un phénotype aggravé par rapport aux souris *Mbd2*^{-/-}, ce qui montre que ces deux protéines interviennent dans une voie fonctionnelle commune, leurs présences paraissant par conséquent plutôt concomitantes dans le complexe Mi2/NuRD.

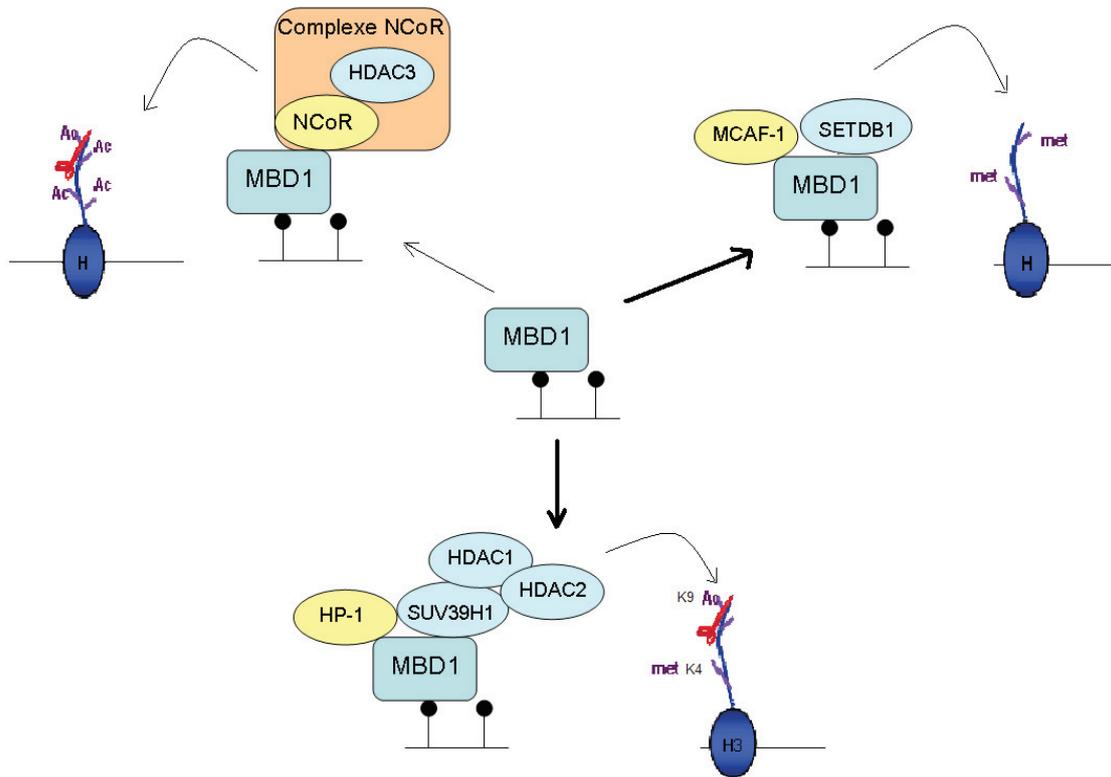


Figure 12 : Le répresseur MBD1 recrute des complexes contenant des HDAC et/ou des HMT lorsqu'il est lié à l'ADN méthylé

Après liaison à l'ADN méthylé, la protéine MBD1 recrute essentiellement des complexes contenant des méthyltransférases d'histones (HMT). Elle interagit avec SETDB1 et le co-répresseur MCAF-1 ce qui permet de méthyler les lysines ou arginines des histones. Elle peut également lier SUV39H1 couplée à HP-1, l'HMT pouvant à son tour recruter les HDAC1 et 2, ce qui permet une désacétylation couplée à une méthylation. Dans quelques cas, la protéine MBD1 peut interagir avec le complexe N-CoR contenant HDAC3 pour désacétyler les histones.

- **MBD1 et MBD4 joueraient un faible rôle dans la répression transcriptionnelle dépendante des HDAC**

Si la protéine MBD1 a une fonction répressive dépendante des désacétylases d'histones (Ng, Jeppesen et al. 2000) puisqu'elle est sensible à l'inhibiteur des HDAC, la trichostatine A, elle diffère des autres protéines à MBD puisqu'elle n'interagit pas avec les désacétylases d'histones 1 et 2. En revanche, une étude montre par co-transfections, CHIP et interférence à l'ARN, que PML-RAR α recrute une DNMT pour méthyler un site et MBD1 qui, après fixation à l'ADN méthylylé *de novo*, pourra induire une répression transcriptionnelle dépendante de la présence de HDAC3 et N-CoR (Villa, Morey et al. 2006) dans des leucémies promyélocyaires aiguës (**Figure 12**).

Enfin, une seule étude fait référence à l'implication de MBD4 dans la répression transcriptionnelle (Kondo, Gu et al. 2005). Des expériences d'immunoprécipitations de chromatine et protéiques révèlent que cette protéine à MBD lie directement les promoteurs méthylylés des gènes *p16^{INK4A}* et *hMLH1* et qu'elle recrute le co-répresseur Sin3A, tout comme MBD2 et MeCP2, et l'HDAC1 pour induire leur inhibition transcriptionnelle.

II.C.1.a.ii Les protéines à MBD initient une répression dépendante de la méthylation de l'ADN en recrutant des méthyltransférases d'histones

Il a été montré qu'en présence de Trichostatine A (TSA), la transcription initialement inhibée par les protéines à MBD n'est pas complètement rétablie (Nan, Ng et al. 1998; Yu, Thiesen et al. 2000). Aussi d'autres mécanismes indépendants des HDAC semblent intervenir dans la répression. Ainsi, les protéines à MBD peuvent interagir avec des complexes contenant des méthyltransférases d'histones (HMT).

- **MeCP2, coREST et HP1 associé à la méthyltransférase SUV39H1**

Des expériences d'immunoprécipitations de chromatine (ChIP) dans des fibroblastes de rat ont montré que le complexe répresseur REST/NRSF contenant co-REST, MeCP2 et

HP-1 sont présents au niveau du promoteur méthylé de *NaCh type II*, un gène exprimé spécifiquement dans les neurones (Lunyak, Burgess et al. 2002). Des expériences de CHIP à partir des fractions immunoprécipitées d'autres CHIP ont révélé que MeCP2 interagit directement avec co-REST, précédemment décrit comme un co-répresseur recruté par REST/NRSF (Andres, Burger et al. 1999) et avec l'HMT SUV39H1 (suppressor of variegation 39H1) qui en diméthylant la lysine 9 de l'histone H3 favorise la liaison HP-1 sur ce site (Lachner, O'Carroll et al. 2001). Tout ce complexe protéique participe à la répression transcriptionnelle de ce gène. Il est à noter que SUV39H1 et HP-1 peuvent aussi participer à l'inhibition transcriptionnelle d'autres gènes comme celui codant pour la cycline E suite à leur recrutement par la protéine Rb (Nielsen, Schneider et al. 2001).

Ces études montrent donc que les protéines à MBD font le lien entre la méthylation de l'ADN et celle des histones. D'autres travaux ont également mis en évidence une corrélation entre l'activité HMT et la présence de MeCP2.

Ainsi, des expériences de purifications d'extraits nucléaires à l'aide d'une protéine MeCP2 complète ou tronquée et couplée à la GST ainsi que des co-immunoprécipitations ont permis de déterminer que la protéine MeCP2 est associée à une activité méthyltransférase d'histones sur H3 *in vitro* et *in vivo* et que cette interaction avec une HMT fait intervenir les domaines MBD et TRD de MeCP2. Des immunoprécipitations de chromatine dans un système inductible Tet-Off pour MeCP2 ont révélé que MeCP2 se lie à une région méthylée du gène *H19* et que la lysine 9 de l'histone H3 est la cible d'une HMT non identifiée (Fuks, Hurd et al. 2003).

- **MBD1 peut interagir avec différentes HMT**

Le mode de répression le plus fréquemment initié par MBD1 apparaît être dépendant d'une activité HMT.

Ainsi, une première étude montre que MBD1 est associée à MCAF1 et qu'une répression transcriptionnelle non dépendante d'une désacétylation d'histones a lieu (Fujita, Takebayashi et al. 1999). Plus tard, un autre groupe va mettre en évidence par co-immunoprécipitations que MBD1 recrute suite à sa liaison au promoteur méthylé d'une part la

HMT SETDB1 qui triméthyle la lysine 9 de H3 et d'autre part MCAF-1 spécifiquement lorsque les cellules sont en phase S (Sarraf and Stancheva 2004). Ce complexe permet alors une répression transcriptionnelle de *p53BP2* dans les cellules HeLa. Par la suite, deux autres équipes établissent que MCAF-1 se lie à MBD1 lorsque celle-ci subit une SUMOylation sur ces lysines (Uchimura, Ichimura et al. 2006; Sekiyama, Ikegami et al. 2008).

De plus, MBD1 interagit, comme MeCP2, avec l'HMT Suv39h1 associée à HP1 (Fujita, Watanabe et al. 2003). En effet, des expériences d'immunofluorescence dans les cellules HeLa montrent que toutes ces protéines co-localisent au niveau d'une hétérochromatine riche en CpG méthylés et contenant des histones H3 méthylées sur la lysine 9. Des expériences de CHIP suivies de PCR ont révélé que MBD1, Suv39H1 et HP1, contrairement à l'ARN polymérase II, sont situées sur le promoteur méthylé de *p16* dans les cellules murines NCI-H1299. Des chromatographies d'affinité « GST-pull down » ont révélé une interaction directe entre MBD1 et Suv39h1 via le domaine MBD. Des mesures d'activité transcriptionnelle *in vitro* via un système gène rapporteur *luciférase* mettent en évidence un rôle de Suv39h1 dans la répression initiée par MBD1. Par contre, d'autres expériences GST-pull down analysées par Western Blot montrent une association des HDAC1 et 2 via leur interaction avec l'HMT Suv39h1.

Aussi, la protéine MBD1 peut enclencher une répression transcriptionnelle en recrutant un complexe HMT qui éventuellement interagit avec des HDAC. Cette combinaison de HMT/HDAC paraît être retrouvée pour la régulation d'autres cibles comme miR-184 car une désacétylation de H3K9 et une méthylation de H3K4 sont observées (Liu, Teng et al.) (Figure 12).

II.C.1.a.iii De multiples modifications de la structure chromatinienne interviennent dans la répression dépendante de la méthylation de l'ADN et initiée par les protéines à MBD

Les protéines à MBD couplent donc la méthylation de l'ADN à la compaction chromatinienne par le recrutement de complexes HDAC et HMT qui catalysent la désacétylation et méthylation des histones respectivement. Ces modifications altèrent

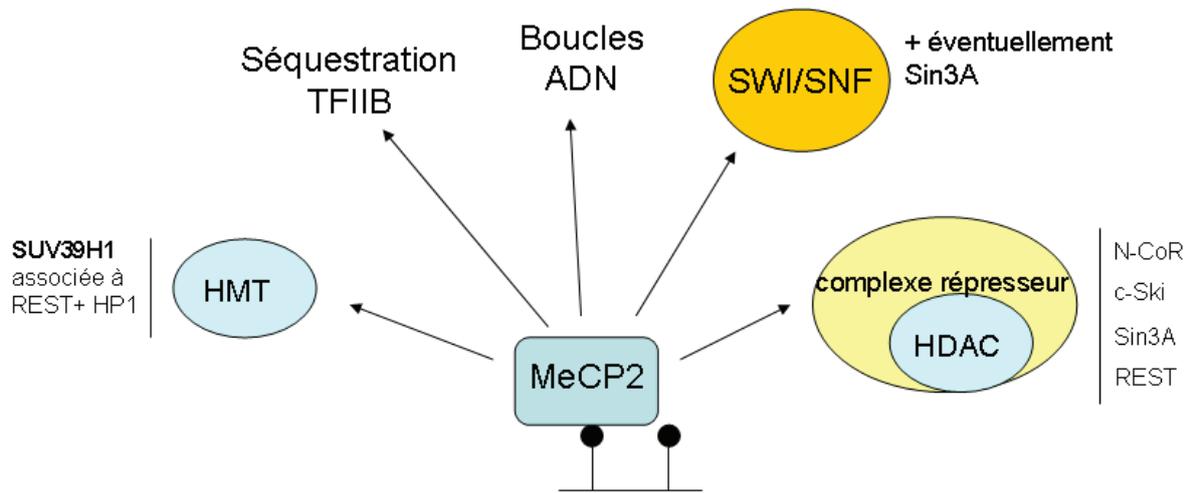


Figure 13 : Répression transcriptionnelle dépendante de la méthylation de l'ADN initiée par la protéine MeCP2.

l'affinité des histones pour l'ADN et la chromatine se retrouve alors condensée et transcriptionnellement inactive.

II.C.1.b La protéine MeCP2 initie d'autres mécanismes dépendant de la méthylation pour remodeler la chromatine

Nous avons vu que la protéine MeCP2 peut initier des changements de conformation de la chromatine dépendant de la méthylation de l'ADN en recrutant des HMT et HDAC.

Il apparaît aussi qu'elle peut interagir avec des complexes remodelant la chromatine sans faire intervenir initialement de HDAC ou HMT ou par la formation de boucles d'ADN. Le recrutement de HDAC peut venir renforcer le dispositif de verrouillage mis en place (**Figure 13**).

II.C.1.b.i MeCP2 et le complexe SWI/SNF remodelant la chromatine

Les complexes SWI/SNF, formés de translocases capables de remodeler la chromatine en hydrolysant l'ATP, modifiant ainsi la structure nucléosomique, peuvent être impliqués dans l'activation ou la répression transcriptionnelle. Des expériences de co-sédimentation, de co-précipitation, de chromatographie d'affinité et de marquages par immunofluorescence ont montré que MeCP2 interagit avec un complexe SWI/SNF via son interaction avec l'unité catalytique Brahma (Brm) et d'autres composants et qu'elle forme un complexe co-répresseur avec Sin3A, HDAC1 et HDAC2 (Jones, Veenstra et al. 1998; Nan, Ng et al. 1998; Harikrishnan, Chow et al. 2005). Ces associations permettent *in vivo* de réprimer les gènes *MDR1*, *THBS1* dans des lignées cancéreuses et *FMRI* dans des cellules issues de patients ayant le syndrome de l'X fragile, après fixation de MeCP2 et Brm à leurs promoteurs méthylés, d'après des « ChIP re-ChIP » (la fraction chromatinienne immunoprécipitée par un premier anticorps sert de chromatine de départ pour une seconde immunoprécipitation), des interférences à l'ARN et l'utilisation d'agents déméthylant ou désacétylant. Cependant, ces résultats n'ont pas pu être reproduits par Hu et ses collaborateurs (Hu, Nan et al. 2006) qui suggèrent notamment que l'autre équipe avait utilisé des anticorps peu spécifiques. Or Harikrishnan et ses collaborateurs ont à leur tour répondu à ces critiques dans le même journal

et les données récemment obtenues par une troisième équipe appuient et corroborent leur première étude (Vecsler, Simon et al.)

II.C.1.b.ii MeCP2 peut former des boucles d'ADN

La protéine MeCP2 peut assurer la formation de boucle d'ADN afin d'isoler des gènes de la machinerie transcriptionnelle. En effet, à partir de cerveaux issus de souris contrôles ou invalidées pour *Mecp2*, des expériences de capture de conformation chromatinienne (3C) ont montré que le gène *Dlx5* est réprimé suite à l'établissement d'une boucle par MeCP2, fixée à des sites méthylés adjacents, au niveau du locus formé par *Dlx5* avec un CGI non méthylé et *Dlx6* (Horike, Cai et al. 2005). La purification par ultracentrifugation des fragments immunoprécipités par CHIP révèle que cette protéine à Mbd recrute à ce locus Hdac1 et qu'elle est associée à une diméthylation de la lysine 9 de l'histone H3, rendant la chromatine transcriptionnellement inactive. Chez les patients atteints du syndrome de Rett, des pertes de l'empreinte parentale paternelle de DLX5 sont retrouvées, suggérant que la conformation secondaire de la chromatine induite par MeCP2 n'est plus assurée. Des études *in vitro* et *in vivo* viennent confirmer l'implication de MeCP2 dans la formation de cette boucle à ce locus (Nikitina, Ghosh et al. 2007) (Yasui, Peddada et al. 2007) alors qu'au même moment un seul travail l'infirme (Schule, Li et al. 2007).

Il est à noter que la protéine MBD2 peut aussi se trouver au niveau de boucles de chromatine (Tiwari, McGarvey et al. 2008) mais elle n'initie pas la formation de cette conformation présente au locus *GATA-4*. Elle paraît plutôt remplacer la protéine PcG2, responsable de ces boucles, dans une zone hyperméthylée pour préserver la conformation chromatinienne.

II.C.1.b.iii MeCP2 peut séquestrer des facteurs appartenant à la machinerie transcriptionnelle

Le facteur de transcription TFIIB appartient à la machinerie mise en place pour l'activation de la transcription. Des expériences de chromatographie d'affinité « GST- pull down » a révélé *in vitro* que la protéine MeCP2 interagit via son domaine TRD avec ce facteur. La séquestration de TFIIB altère alors la pré-initiation de la transcription. De plus, un

système de transcription *in vitro* a montré que MeCP2 via son domaine TRD participe à une répression transcriptionnelle, suite à sa fixation sur des sites méthylés, qui n'est pas levée après l'ajout de TFIIB ou d'agents inhibiteurs des HDAC (Kaludov and Wolffe 2000). MeCP2 peut donc jouer un rôle direct et indirect dans la répression transcriptionnelle, indépendamment de l'acétylation des histones, ce qui reste à confirmer *in vivo*. D'autre part, comme il l'a été évoqué précédemment, N-CoR interagit avec MeCP2 et d'après une autre étude menée *in vivo* avec TFIIB et TAF_{II}32, un autre membre du complexe d'initiation de la transcription, les empêchant de remplir leurs rôles (Muscat, Burke et al. 1998). Il reste à définir une éventuelle liaison entre ces événements.

II.C.1.c La répression dépendante de la méthylation serait favorisée au sein de certaines séquences consensus

Les expériences de retard sur gel menées par l'équipe de Bird (Hendrich and Bird 1998) montraient que les protéines à MBD liaient les versions méthylées de différentes sondes ne présentant aucunes séquences nucléotidiques communes. Ceci laissait envisager que la liaison des protéines à MBD, via leur domaine MBD, s'effectuait sur les dinucléotides CpG méthylés de façon séquence-indépendante. Toutefois, Fraga et ses collaborateurs (Fraga, Ballestar et al. 2003) ont nuancé cette observation suite à leur mesure de l'affinité des MBD pour différentes séquences par une technique d'électromobilité sur capillaires. La densité, le nombre de CpG ainsi que les séquences adjacentes aux séquences méthylées pourraient influencer la fixation naturelle des protéines à MBD.

La première spécificité de séquence a été rapporté pour MeCP2 par l'équipe du Pr. A. Bird (Klose, Sarraf et al. 2005). En effet, la génération de fragments ADN contenant en leur centre un site CCGG méthylé *in vitro* et de part et d'autres différentes séquences a permis de montrer que, contrairement à MBD2, MeCP2 présente une forte affinité pour ceux contenant un ou plusieurs CpG méthylé(s) et des motifs de 4 A/T adjacents. De plus, ces séquences A/T sont retrouvées au niveau des gènes *Bdnf* et *Dlx6* pour lesquels la présence de MeCP2 sur leurs promoteurs avait déjà été identifiée (Martinowich, Hattori et al. 2003; Zhou, Hong et al. 2006). Ceci semble donc confirmer que les séquences A/T jouent un rôle essentiel dans la liaison de MeCP2 même si MeCP2 n'interagit pas avec tous ces sites potentiels.

Plus récemment, une spécificité de séquence pour MBD1 a été identifiée *in vitro* suite à des expériences de retard sur gel avec des protéines MBD1 étiquetées avec la GST et incubées avec différentes séquences mutées autour d'un CpG méthylé ou non ainsi qu'à l'utilisation de banques d'ADN contenant ou non des CpG méthylés (methyl-SELEX) (Clouaire, de Las Heras et al.). La fixation de MBD1 par son domaine MBD sur un CpG méthylé serait favorisée par la présence d'autres CpG méthylés successifs ou par sa localisation dans un motif (TouTG)CGCA. Ces séquences, retrouvées *in vivo* au sein des promoteurs des gènes *HBA*, *NGFR* et *RND2*, ne sont pas liées par MBD2. En revanche, le domaine CXXC3 ne présente aucune affinité de liaison particulière et ne participe pas au recrutement de MBD1 au niveau de ces cibles privilégiées.

Ainsi, certaines protéines à MBD semblent sensibles aux séquences nucléotidiques présentes autour des sites méthylés.

II.C.2 Des répresseurs également recrutés par des facteurs de ciblage liés à l'ADN

Des facteurs de transcription possédant des sites de liaison spécifiques à l'ADN, dépendant ou non de la présence de méthylation selon les cas, peuvent s'associer aux protéines à MBD pour recruter via leur intermédiaire des complexes enzymatiques impliqués dans le remodelage de la chromatine et ainsi permettre une répression transcriptionnelle. Les protéines à MBD agiraient alors comme des co-répresseurs.

II.C.2.a Quand le complexe Mi2/NuRD est recruté via MBD2

La protéine MiZF (MBD2-interacting zinc finger protein) a été identifiée par double hybride comme co-facteur de la protéine MBD2 (Sekimata, Takahashi et al. 2001), caractéristique confirmée *in vivo*. Des expériences de gènes rapporteurs et l'utilisation d'agent inhibiteur des HDAC ont montré que MiZF recrute MBD2 ainsi que des complexes contenant des HDAC pour induire une répression transcriptionnelle. De plus, il apparaît que MiZF reconnaît une séquence nucléotidique consensus indépendamment de l'état de méthylation

(Sekimata and Homma 2004) guidant ainsi le recrutement de MBD2 sur des sites non méthylés et/ou méthylés qu'elle ne ciblerait pas seule.

Le complexe Mi2/NuRD peut aussi être recruté par la protéine GATA-1 via la protéine FOG-1 qui sert de pont protéique, comme le montre des expériences de spectrométrie de masse et des transfections transitoires associées à des immunoprécipitations (Rodriguez, Bonte et al. 2005). Un système de gène rapporteur sous contrôle d'un promoteur contenant des sites de fixation de la protéine GATA-1 a permis de déterminer que le complexe GATA-1/FOG-1/MeCP1 se lie via GATA-1 sur le promoteur, ce qui a pour conséquence de réprimer le gène. Ces résultats ont été confirmés par des expériences de CHIP au niveau du gène *GATA-2*, et par une autre équipe pour *c-kit* (Hong, Nakazawa et al. 2005).

II.C.2.b MeCP2 et PU.1

Il a été montré par chromatographie sur colonne d'affinité que PU.1, un facteur de transcription hématopoétique aussi appelé SPI-1, peut recruter la protéine MeCP2 par le domaine TRD et par ChIP que ce complexe se fixe via PU.1 spécifiquement sur le gène de la β -globine. Des expériences de gènes rapporteurs ainsi que l'utilisation d'agent inhibiteur de HDAC ont montré que la protéine MeCP2 met en jeu à son tour le complexe mSin3A-HDAC pour une répression de la transcription (Suzuki, Yamada et al. 2003).

II.C.2.c MBD2, MeCP2 et ARN interférants

Les ARN interférants, qui s'hybrident avec leurs cibles complémentaires génomiques semblent aussi être des partenaires des protéines à MBD. En effet, les protéines MBD2 et MeCP2 présentent une région riche en arginines et en acide glutamique (région RG) donc sont susceptibles d'interagir avec l'ARN. Une incubation de tous les types d'ARN avec les protéines à MBD purifiées, entières ou tronquées, a permis d'établir qu'un complexe protéine-ARN ou RNP se forme *in vitro* entre les ARN messagers ou interférants et les protéines MBD2 ou MeCP2 via leur région RG (Jeffery and Nakielny 2004). De plus, il y a compétition entre les ARN interférants et des fragments d'ADN méthylé pour leur fixation à ces protéines

à MBD, bien que les domaines de liaison diffèrent. Ceci suggère que les ARN peuvent guider la fixation des protéines à MBD ou réguler négativement leur action.

II.C.3 Une activité répressive modulée

Les protéines à MBD interviennent toutes dans la répression transcriptionnelle. Cependant, ces activités peuvent être modulées par l'intervention d'autres facteurs, des modifications post-traductionnelles ou être impliquées dans d'autres mécanismes de régulation transcriptionnelle comme l'épissage.

II.C.3.a MBD2 et MeCP2 recrutent différents partenaires pour intervenir directement dans l'activation transcriptionnelle

MBD2 paraît être impliquée dans l'activation de la transcription, pas par une activité déméthylase comme tout d'abord suggéré (Detich, Theberge et al. 2002) mais plutôt par le recrutement de facteurs co-activateurs ou en devenant elle-même une protéine co-activatrice.

Par un système de gène rapporteur, il a été mis en évidence que la protéine MBD2 peut recruter la GTPase MBDin (MBD2 interacting protein), ce qui a pour conséquence une activation transcriptionnelle, sans altération de la méthylation sur le site lié par MBD2 (Lembo, Pero et al. 2003). Ces données sont également retrouvées lors du recrutement de TACC3 (Transforming Acid Coiled Coil) (Angrisano, Lembo et al. 2006). Cette étude laisse également apparaître que le complexe MBD2/TACC3 est associé à une acétyltransférase d'histones (pCAF), permettant probablement de favoriser la décompaction chromatinienne. Aussi, la protéine MBD2 paraît capable d'interagir avec des co-activateurs.

Des expériences de co-immunoprécipitations et de chromatographie d'affinité « GST-pull down » ont révélé que MBD2a, via sa liaison avec RHA (RNA helicase A), interagit avec le complexe CREB (Fujita, Fujii et al. 2003), contrairement à MBD2b. Cette interaction n'induit pas le recrutement de HDAC1. En revanche, d'après un système de gène rapporteur, cette association a un effet synergique sur la transcription initiée par la liaison de CREB phosphorylée sur un motif CRE. L'expression exogène de MBD2a, RHA dans des cellules murines modifiées contenant le gène de la somatostatine du rat a permis de confirmer d'une

part l'association entre MBD2a et RHA par CHIP et d'autre part que cette liaison sur une région promotrice entraînait une activation transcriptionnelle par RT-PCR semi-quantitative. A ce moment-là, MBD2 agit comme co-activateur dans la transcription génique.

Des analyses transcriptomiques d'hypothalamus de souris, invalidées ou non pour *Mecp2*, ont permis de mettre en évidence que *Mecp2* induit en majorité l'activation de cibles. Ce mécanisme est direct puisque des CHIP anti-*Mecp2* ont montré la fixation de cette protéine sur leurs promoteurs peu méthylés, d'après des modifications au bisulfite de sodium et séquençage. *Mecp2* paraît recruter *Creb1* comme co-activateur, au moins pour certains gènes (Chahrour, Jung et al. 2008). Ces données sont en accord avec celles issues d'analyses « CHIP-on-chip » où il apparaît que MeCP2 est souvent détectée sur des promoteurs où l'ARN polymérase II est aussi fixée (Yasui, Peddada et al. 2007).

Ainsi, selon les facteurs que les protéines MBD2 ou MeCP2 recrutent au niveau des sites méthylés, elles semblent initier soit la répression soit l'activation transcriptionnelle.

II.C.3.b MBD2 et MeCP2 ont aussi des rôles indirects dans l'activation transcriptionnelle

II.C.3.b.i MBD2 entre en compétition avec des activateurs transcriptionnels

Comme il l'a été évoqué précédemment, les analogues de MBD3, MBD3L1 ou MBD3L2, peuvent remplacer les protéines MBD3 ou MBD2 respectivement dans le complexe Mi-2/NuRD, ce qui implique une compétition de liaison entre ces protéines (Jiang, Jin et al. 2004; Jin, Jiang et al. 2005). Ces analogues mimeraient l'action des protéines MBD2 et MBD3 puisque ces changements ne semblent pas empêcher le complexe répresseur de participer à une inhibition transcriptionnelle dépendante de la méthylation de l'ADN. Par contre, certaines compétitions entre les protéines à MBD et d'autres facteurs de transcription peuvent mutuellement empêcher leur action sur la régulation génique.

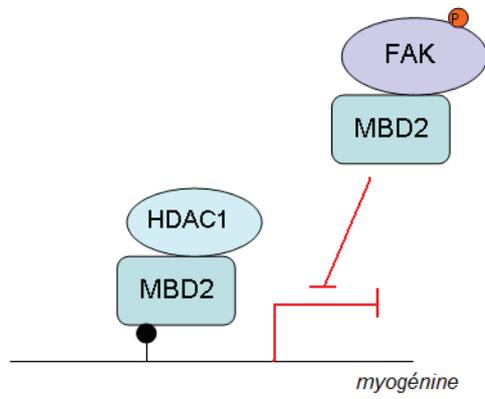


Figure 14 : Régulation transcriptionnelle de la myogénine par compétition de liaison entre HDAC1 et FAK phosphorylée pour la protéine MBD2

- **Compétitions de liaison à MBD2 entre HDAC1 et FAK**

Par des constructions de mutants, des chromatographies d'affinité « GST-pull-down » et des co-immunoprécipitations, il a été montré que MBD2a interagit avec les protéines PYK2 et FAK (Focal adhesion kinase) via leur domaine FAT (Luo, Zhang et al. 2009). FAK joue un rôle important dans les adhésions des cellules à la matrice extracellulaire et a une localisation cytoplasmique. Cependant, cette protéine peut devenir nucléaire après phosphorylation et liaison avec MBD2 dans les myotubes et les myoblastes. Des ChIP montrent que MBD2 lie le promoteur méthylé de la myogénine mais cette interaction est diminuée suite à l'expression exogène de FAK. Des co-transfections suivies d'immuno-précipitations ont mis en évidence une compétition de liaison à MBD2 entre FAK et HDAC1 (**Figure 14**).

Aussi la formation du complexe MBD2-FAK réduirait la fixation de MBD2 à l'ADN méthylé ainsi que le recrutement d'enzymes impliquées dans la compaction de la chromatine.

- **Compétitions de fixation à l'ADN**

Il existe une compétition de fixation sur le promoteur du gène *Il4* entre le répresseur MBD2 et l'activateur GATA-3 (Hutchins, Mullen et al. 2002). Or *IL4* est impliquée dans la différenciation des lymphocytes T, ce qui pourrait expliquer pourquoi suite à l'inactivation de *Mbd2* chez la souris, une altération de ce mécanisme est observée.

La protéine MiZF, recrutant la protéine MBD2 afin d'induire une répression transcriptionnelle (Sekimata, Takahashi et al. 2001; Sekimata and Homma 2004) peut également faire concurrence à l'activateur ATF au niveau du gène du rétinoblastome (*Rb*).

Ainsi, la protéine MBD2 peut entrer en compétition directe ou indirecte avec des activateurs transcriptionnels, ce qui a pour effet de moduler son activité répressive.

Elle paraît également capable de recruter des co-activateurs ou d'agir comme un co-activateur pour jouer un rôle antagoniste à celui de répresseur.

II.C.3.b.ii MeCP2 intervient dans l'épissage alternatif

La protéine MeCP2 après phosphorylation par CDKL5 peut intervenir dans l'épissage alternatif (Young, Hong et al. 2005; Ricciardi, Kilstrup-Nielsen et al. 2009). En effet, des co-immunoprécipitations ont montré que MeCP2 interagit avec la protéine de liaison aux ADN/ARN YB-1 (Y-box-binding protein 1) et un modèle de mini-gènes rapporteurs possédant plusieurs sites d'épissages a permis de définir que ce complexe pouvait prendre en charge l'ARN précurseur transcrit du gène rapporteur et d'en orienter l'épissage alternatif.

Aussi, la protéine MeCP2 peut intervenir dans la régulation génique au niveau du gène ou de l'ARNm.

II.C.3.c Les modifications post-traductionnelles des protéines à MBD modulent leur activité répressive

Comme de nombreuses protéines impliquées dans la régulation transcriptionnelle, certaines protéines à MBD peuvent subir des modifications post-traductionnelles qui vont permettre de moduler leur activité de répresseur. Il apparaît que ces modifications, détaillées dans ce paragraphe, n'influencent pas la localisation des protéines à MBD au niveau de l'hétérochromatine.

II.C.3.c.i Phosphorylation/déphosphorylation de la protéine MeCP2

Il a été vu précédemment que, quand les neurones sont quiescents, seule la protéine MeCP2 occupe le promoteur du gène tissu-spécifique *Bdnf* (Zhou, Hong et al. 2006) codant pour une protéine responsable de la survie et croissance neuronale. La phosphorylation de MeCP2 au niveau de la sérine 421 par CDKL5 dans un contexte voltage-dépendant entraîne le détachement de MeCP2 couplé à HDAC1 et Sin3A et paraît favoriser l'acétylation de l'histone H3, avec pour conséquence finale la levée de l'inhibition transcriptionnelle (Chen, Chang et al. 2003).

A l'inverse, il apparaît que la phosphorylation de la protéine MeCP2 au niveau d'une autre sérine, la S80, favorise quant à elle l'association de cette protéine aux promoteurs de

différents gènes *Pomc*, *Gtl2*, *Rad3d*, *Vamp3* et *Igsf4b* dans des cellules neuronales de souris (Tao, Hu et al. 2009). Plus tard, la kinase HIPK2 (homeodomain-interacting protein kinase 2) sera identifiée comme responsable de la phosphorylation sur la sérine 80 *in vitro* et *in vivo* (Bracaglia, Conca et al. 2009; Ricciardi, Kilstrup-Nielsen et al. 2009). La protéine MeCP2 phosphorylée aurait alors un rôle pro-apoptotique dans le cerveau (Bracaglia, Conca et al. 2009). L'invalidation du gène codant *hipk2* (Isono, Nemoto et al. 2006) semble également provoquer des troubles moteurs similaires à ceux visualisés après invalidation du gène *MeCP2* ou suite à la déphosphorylation ou à la mutation de cette sérine en arginine, confirmant l'existence d'un lien étroit entre la phosphorylation en S80 de MeCP2 et HIPK2.

Ainsi, la protéine MeCP2 semble régulée de manière dynamique par la phosphorylation et déphosphorylation de multiples sérines selon l'activité neuronale.

II.C.3.c.ii Méthylation des arginines de la protéine MBD2

La protéine MBD2 peut subir des modifications post-traductionnelles différentes de MeCP2. Ainsi, la protéine MBD2 co-précipite avec les méthyltransférases d'arginines PRMT1 et PRMT5 qui catalysent *in vitro* et *in vivo* la méthylation d'arginines présentes au niveau du domaine RG de MBD2 (Le Guezennec, Vermeulen et al. 2006; Tan and Nakielny 2006). Ceci a pour conséquence une diminution de l'affinité de MBD2 pour les CpG méthylés ainsi que pour le complexe Mi2/NuRD notamment via son interaction avec les désacétylases d'histones HDAC1 et 2. L'utilisation d'un système de gène rapporteur luciférase, sous contrôle d'un promoteur situé en aval de motifs Gal4 ciblés par la protéine chimérique GAL4-MBD2 (domaine de liaison à l'ADN de GAL4 fusionné avec différents mutants de MBD2) a permis de montrer que la protéine MBD2 ne remplit plus son rôle de répresseur suite à cette modification dans le domaine RG.

II.C.3.c.iii SUMOylation de la protéine MBD1

La SUMOylation (Small Ubiquitin-like Modifier) concerne de nombreuses protéines impliquées dans la régulation transcriptionnelle. Contrairement à l'ubiquitinylation, elle ne provoque pas la dégradation de protéines. Elle touche les résidus lysines présents dans un

motif peptidique consensus. Des expériences de chromatographie d'affinité « GST pull-down » et d'interférences à l'ARN ont permis de mettre en évidence une modification de MBD1 par SUMO1, 2 et 3 *in vivo* et *in vitro*, favorisant l'interaction entre MBD1 et MCAF-1 via ces modifications présentes dans le domaine TRD, avec une incidence sur la triméthylation de H3K9 (Uchimura, Ichimura et al. 2006). Par la suite, une autre équipe a permis de préciser que les deux ligases E3, PIAS1 et PIAS3 (Protein Inhibitors of Activated STAT) contribuent à la conjugaison de SUMO1 sur MBD1, au niveau des lysines 450 et 489 dans un contexte VKQE dans les cellules HeLa (Lyst, Nan et al. 2006). De plus, même si MBD1 peut encore lier l'ADN méthylé et MCAF-1 via sa modification SUMO-3, elle ne peut plus recruter la méthyltransférase d'histones SETDB1 donc H3K9 n'est pas triméthylée et la répression de la transcription du gène *p53PB2* n'a plus lieu (Sekiyama, Ikegami et al. 2008).

II.C.4 Les protéines à MBD servent d'intermédiaires entre la méthylation de l'ADN et les changements chromatinien conduisant à la répression transcriptionnelle

S'il avait été découvert une corrélation entre la méthylation de l'ADN et la désacétylation des histones, il aura fallu attendre la découverte de MeCP2 comme protéine qui lie l'ADN méthylé (Meehan, Lewis et al. 1992) et qui recrute un complexe enzymatique avec une activité désacétylase (Jones, Veenstra et al. 1998; Nan, Ng et al. 1998) pour établir un lien direct entre méthylation de l'ADN et changements chromatinien. Ces 2 mécanismes épigénétiques agissent en synergie notamment dans la répression transcriptionnelle de gènes dans les cancers (Cameron, Bachman et al. 1999). Puis suivront les découvertes des autres protéines à MBD qui peuvent remplir ce rôle ou agir via d'autres voies pour participer à l'inhibition transcriptionnelle.

Aussi, identifier les gènes ciblés par les protéines à MBD pourrait permettre de comprendre, au moins en partie, comment l'altération des profils de méthylation de l'ADN dans les cancers participe aux changements d'expression génique.

II.D Balance entre spécificités de cibles et tissulaires et redondances

II.D.1 Spécificités d'un locus pour une protéine à MBD

II.D.1.a Mises en évidence de spécificités de cibles et tissulaires

De nombreuses études démontrent qu'une seule protéine à MBD peut fixer un locus.

En effet, mon équipe a montré que la protéine MBD2 lie spécifiquement les promoteurs méthylés *BRCA1-NBR2* (Auriol, Billard et al. 2005), *pS2/TFE1* (Chatagnon, Ballestar et al.), *hTERT* (Chatagnon, Bougel et al. 2009) et dans des cellules déficientes en MBD2, aucune protéine à MBD ne vient la remplacer sur ce site.

De même, des gènes suppresseurs de tumeurs voient, dans les cancers, leurs promoteurs hyperméthylés occupés par une unique protéine à MBD. Ainsi, des immunoprécipitations de chromatine (ChIP) (Bakker, Lin et al. 2002) ont révélé que, dans les cancers hépato-cellulaires humains, le promoteur de *GSTP1* (glutathione S-transférase) est hyperméthylé et que seule MBD2 vient se fixer afin d'initier la répression de ce gène. De nombreux autres exemples sont donnés par l'étude de Lopez-Serra (Lopez-Serra, Ballestar et al. 2006). Parmi ceux-ci, MBD1 se fixe le promoteur de *p16* dans les cellules H1299, le promoteur de *RASSF1A* dans les cellules MCF7.

Toutes ces données corroborent donc le fait qu'il existe des spécificités de cibles associées à des spécificités de types cellulaires.

D'autres travaux viennent conforter cette hypothèse (Muller, Wischnewski et al.; Ballestar, Paz et al. 2003; Lopez-Serra, Ballestar et al. 2006). En effet, l'occupation de différents promoteurs de gènes suppresseurs de tumeurs par les protéines à MBD a été étudiée dans diverses lignées cellulaires cancéreuses (Ballestar, Paz et al. 2003; Lopez-Serra, Ballestar et al. 2006). A titre d'exemple, il apparaît que le promoteur méthylé du gène *CDH13* est fixé par uniquement par MeCP2 dans les cellules MCF7.

Tous ces résultats indiquent que la distribution des protéines à MBD sur les promoteurs étudiés est dépendante du type cellulaire et que leur présence est largement requise pour verrouiller ces cibles.

II.D.1.b Comment expliquer les spécificités de cibles des protéines à MBD

Ces spécificités de cibles pourraient s'expliquer en partie par l'abondance des protéines à MBD qui s'accompagnerait d'un phénomène de redistribution de ces répresseurs. Ainsi, il a été vu que dans les cellules déficientes en MBD2, la protéine MeCP2 ne se fixe pas au promoteur *BRCA1-NBR2*. Par contre, une sur-expression exogène de *Mecp2* entraîne sa liaison à cette région (Auriol, Billard et al. 2005). Cependant, la concentration protéique n'est pas un critère suffisant pour décrire toutes les spécificités de cibles (Lopez-Serra, Ballestar et al. 2006). Elle pourrait même dans certaines cellules être la marque d'une répression transcriptionnelle systématisée c'est-à-dire non spécifique. En effet, dans les neurones, MeCP2 est très abondante et paraît associée à une répression dépendante des HDAC et localisée au niveau de la plupart des sites méthylés (Skene, Illingworth et al.).

Par contre, il a été évoqué précédemment l'existence de facteurs nucléaires semblant capables de diriger les protéines à MBD et ainsi moduler leur activité répressive. L'environnement dans lequel se trouvent les protéines à MBD paraît donc également influencer leur fixation.

Le recrutement de ces protéines sur des leurs cibles spécifiques méthylées serait dû à la combinaison complexe de spécificités de séquences, d'effets de concentration d'une protéine par rapport à une autre et d'intervention de facteurs nucléaires. De plus, leurs spécificités semblent dépendantes du tissu considéré et de son état par rapport au cancer.

A l'inverse, si MBD2 peut se fixer sur des sites laissés vacants par MeCP2, l'inverse n'est pas toujours valable (Klose, Sarraf et al. 2005). De plus, des ChIP menées sur des cellules cancéreuses (Ballestar, Paz et al. 2003) montrent qu'environ 9% des sites occupés par MeCP2 ne le sont pas par MBD1 et MBD2. Ces données mettent en évidence qu'il existe également un processus de remplacement d'une protéine à MBD par une autre.

II.D.2 Quand plusieurs protéines à MBD occupent le même site

Plusieurs études concluent que plusieurs protéines à MBD peuvent occuper un même site. Ainsi, les protéines MeCP2, MBD1 et MBD2 sont retrouvées pour le promoteur du gène *CDH1* d'où une redondance vis-à-vis de l'occupation de ce site (Bakker, Lin et al. 2002).

De plus, l'équipe du Dr. Esteller (Ballestar, Paz et al. 2003) ont étudié des séquences répétées comme *Sat2* et les promoteurs des gènes soumis à empreinte parentale comme *Igf2*. Leurs résultats indiquent que plusieurs protéines à MBD occupent ces sites.

Des CHIP suivis d'autres CHIP que les protéines MeCP2 et le complexe MeCP1 peuvent se fixer simultanément sur le promoteur du gène à copie unique *SYBL1* (Matarazzo, De Bonis et al. 2007). La protéine Kaiso ainsi qu'un complexe polycombe PRC2 impliqué dans la répression transcriptionnelle semblent également pouvoir occuper le promoteur de ce gène.

Il semble donc que plusieurs mécanismes soient mis en place afin de garantir le verrouillage de certaines cibles, soit par une redondance des protéines à MBD soit par l'action d'autres protéines de liaison à l'ADN ou complexes répresseurs. Ceci suggère qu'il est important que ces cibles demeurent transcriptionnellement inactives. Il est également à noter que peu d'études utilisent la technique de CHIP suivi d'un autre CHIP. Ceci explique pourquoi la présence simultanée ou non de plusieurs protéines à MBD ou d'autres répresseurs n'est pas connue sur de nombreux loci.

II.D.3 Spécificités de cibles et tissulaires se mêlent à un phénomène de redondance

Au sein des mêmes cellules, des sites ciblés par une seule protéine à MBD co-existent avec d'autres loci fixés par plusieurs protéines à MBD. Il semble donc que les protéines à MBD possèdent des spécificités tissulaires et fonctionnelles même si certains phénomènes de redondance sont observés.

Ces résultats sont en accord avec les phénotypes murins observés suite aux invalidations simples ou multiples des gènes codant ces protéines à MBD. Les doubles mutants *Mbd2*^{-/-} et *Mbd3*^{+/-} (Hendrich, Guy et al. 2001) présentent un phénotype aggravé par

rapport aux simples mutants, suggérant une voie fonctionnelle commune entre *Mbd2* et *Mbd3*, c'est-à-dire le complexe Mi2/NuRD. En revanche, une double invalidation de *Mecp2* et *Mbd2* (Guy, Hendrich et al. 2001) ou de *Mecp2* et *Mbd1* ne donne aucune aggravation des phénotypes, laissant supposer que ces protéines ne partagent pas les mêmes fonctions.

Ces données corrélient également par exemple avec des analyses transcriptomiques de souris témoins et déficientes en *Mecp2* (Tudor, Akbarian et al. 2002). En effet, si aucun changement significatif n'est repéré, la sélection de quelques gènes à partir de ces puces a permis de mettre en évidence des différences d'expression faibles mais significatives pour ces gènes. Ces résultats suggèrent donc que la perte de *Mecp2* peut conduire à des changements d'expression génique mais globalement très peu de gènes sont concernés. Ceci reste donc plutôt en faveur d'une redondance avec une spécificité pour quelques gènes. Par ailleurs, l'analyse transcriptomique de cellules HeLa ayant été transfectées par des petits ARN interférants ciblant MBD1 et/ou MBD2 et/ou MeCP2 a révélé que chaque interférence à l'ARN provoquait l'altération d'un nombre différent de gènes mais que dans les doubles ou triples transfections, le nombre de gènes altérés ne correspondait pas à leurs additions (Lopez-Serra, Ballestar et al. 2008) suggérant ainsi qu'il existe un ensemble de cibles spécifiques et non spécifiques pour les protéines à MBD.

II.E Implications des protéines à MBD dans d'autres processus biologiques

Les membres de la famille des protéines à MBD apparaissent impliqués dans divers autres processus biologiques. Ces rôles ont d'abord été supposés suite aux études des souris nullizygotes pour les protéines à MBD qui ont révélé quelques différences phénotypiques. D'autres travaux sont venus par la suite confirmer et compléter ces observations.

II.E.1 MeCP2, MBD2, MBD1 et leurs implications dans le développement du système nerveux

L'invalidation de *Mecp2* (Hendrich, Guy et al. 2001) provoque de sévères symptômes neurofonctionnels assimilables au syndrome de Rett (Amir, Van den Veyver et al. 1999) avec

notamment un retard mental associé à des troubles psychomoteurs. Le lien étroit existant entre *MeCP2* et ses mutations et ce syndrome a été confirmé et donne encore lieu à de nombreuses publications (Wood and Shepherd; Swanberg, Nagarajan et al. 2009). Ces données sont en accord avec celles concernant son implication dans la différenciation des cellules embryonnaires conduisant à la formation du système nerveux (différenciation neurale) (Martin Caballero, Hansen et al. 2009) et particulièrement dans celle des neurones (Macdonald, Verster et al.; Tao, Hu et al. 2009; Tsujimura, Abematsu et al. 2009). *MeCP2* interviendrait aussi dans la maturation neuronale.

La protéine *Mbd2* aurait également un rôle redondant à celui de *Mecp2* dans la différenciation embryonnaire neurale (Martin Caballero, Hansen et al. 2009) et distinct dans la différenciation neuronale (Macdonald, Verster et al.).

La protéine *MBD1* participerait quant à elle à la différenciation neuronale à partir de progéniteurs neuraux chez l'adulte. L'inactivation de *Mbd1* montre que la protéine associée participerait aussi à l'activité hippocampique (Zhao, Ueba et al. 2003).

II.E.2 MeCP2, MBD4, MBD1 et maintien de l'intégrité du génome

Les invalidations de *Mbd1* ont révélé que la protéine *Mbd1* participe à la stabilité chromosomique (Zhao, Ueba et al. 2003) et indirectement à la réparation de l'ADN en ayant plutôt une fonction de réservoir pour MPG, une enzyme qui permet l'excision de bases incorporées à l'ADN par erreur (Watanabe, Ichimura et al. 2003). Elle intervient donc dans le maintien de l'intégrité du génome.

La protéine *MBD4* participe également à cette grande fonction biologique. En effet, les souris nullizygotes *Mbd4*^{-/-} présentent des défauts dans la réparation de l'ADN (Millar, Guy et al. 2002). Cette protéine avait déjà été identifiée comme partenaire de la protéine *MLH-1* impliquée dans ce mécanisme (Bellacosa, Cicchillitti et al. 1999). La protéine *MBD4* protégerait l'ADN des désaminations spontanées ou induites par des drogues (Cortellino, Turner et al. 2003; Sansom, Zabkiewicz et al. 2003) d'où sa fonction dans la réponse apoptotique suites aux dommages de l'ADN (Screaton, Kiessling et al. 2003).

La perte partielle de *MeCP2*, suite à une interférence à l'ARN ciblant l'ARNm *MeCP2* dans des cellules souches mésenchymateuses, induit une sous-expression de gènes impliqués dans la majorité des systèmes de réparation de l'ADN, conduisant à la sénescence par les

voies impliquant P53 et RB. Ces résultats suggèrent donc que la protéine MeCP2 intervient aussi dans l'intégrité du génome en participant à la régulation des systèmes de réparation de l'ADN (Squillaro, Alessio et al.).

II.E.3 Transmission des marques épigénétiques

Certaines protéines à MBD pourraient favoriser la transmission des profils de méthylation de l'ADN. En effet, les protéines MBD2 et MBD3 sont capables de s'hétérodimériser afin de former un complexe présentant une haute affinité pour l'ADN hémi-méthylé. Pendant la phase de réplication de l'ADN, elles viendraient recruter DNMT1 au niveau des foyers de réplication afin que DNMT1 remplisse son activité de maintenance de la méthylation de l'ADN (Tatematsu, Yamazaki et al. 2000). Dnmt1 peut également interagir avec Mecp2, comme le montre des expériences de chromatographie d'affinité, afin d'être guidée vers des sites hémi-méthylés (Kimura and Shiota 2003).

Par contre, MBD1 paraît plutôt impliquée dans la transmission de la méthylation présente sur la lysine 9 de l'histone H3 au niveau de certains loci méthylés comme p53BP2 dans les cellules HeLa. Cette transmission, assurée par l'activité méthyltransférase de SETDB1, est permise par l'association de cette enzyme à MBD1 qui interagit elle-même avec le complexe CAF-1/PCNA/DNMT1 en aval de la fourche de réplication de l'ADN (Sarraf and Stancheva 2004).

II.E.4 Une famille de protéines à MBD qui couvre de nombreux mécanismes cellulaires

Les protéines à MBD interviennent dans de multiples processus avec pour toutes un rôle dans la répression transcriptionnelle. Certaines de ces protéines participent également à d'autres fonctions comme la réparation de l'ADN, la réponse apoptotique lorsque les dommages de l'ADN sont trop importants, le cycle cellulaire. Toutes ces données suggèrent qu'elles occupent un rôle important dans la cancérogenèse.

II.F La protéine MBD2 et la cancérogenèse

Toutes les protéines à MBD paraissent impliquer dans la cancérogenèse. Cependant, seule la contribution de MBD2 dans ce processus sera abordée au cours de cette partie, en outre parce qu'elle paraît être la protéine à MBD réprimant le plus de gènes dans les cancers notamment d'après diverses études par approches globales de l'équipe du Dr Esteller (Ballestar, Paz et al. 2003; Lopez-Serra, Ballestar et al. 2006; Lopez-Serra, Ballestar et al. 2008).

II.F.1 Une combinaison de dérégulations du message et du lecteur

II.F.1.a La protéine MBD2 présente une expression altérée dans les cancers

Etant donné que dans les cancers, les profils de méthylation de l'ADN sont altérés, il a été supposé que les taux des protéines traduisant ce signal soient également modifiés. De nombreuses études se sont attachées à déterminer par PCR quantitatives ou semi-quantitatives le taux d'ARNm *MBD2*, corrélé à la concentration de protéines (Billard, Magdinier et al. 2002).

Il apparaît que le taux d'ARNm *MBD2* dans des tumeurs primaires de poumon (de type NSCLC pour non small cell lung cancer), ovaire, colon et sein est diminué de 20 à 80% par rapport aux contrôles sains (Muller-Tidow, Kugler et al. 2001), de même dans les cancers colorectaux (Kanai, Ushijima et al. 1999), stomacaux ou gastriques (Oue, Kuraoka et al. 2001; Fang, Cheng et al. 2004). De plus, si une augmentation dans les tumeurs de l'endomètre est retrouvée (Muller-Tidow, Kugler et al. 2001), aucune différence significative n'a été observé dans les leucémies myéloïdes.

Ces résultats sont à nuancer car il n'apparaît pas de différences significatives pour l'ARNm *MBD2* entre les tissus cancéreux et les tumeurs bénignes mammaires (Muller, Fiegl et al. 2003), colorectaux (Schmidt, Sedivy et al. 2007) ou du poumon (de type NSCLC) (Sato, Horio et al. 2002) et des augmentations sont même retrouvées dans les cancers du sein (Billard, Magdinier et al. 2002; Tryndyak, Kovalchuk et al. 2006) ou du poumon (de type SCLC) (Sato, Horio et al. 2002). Ces différentes observations peuvent s'expliquer par le choix du gène utilisé pour normaliser les résultats et par le fait qu'il existe une variabilité importante

gènes	cancers associées à cette hyperméthylation	références bibliographiques
<i>BRCA1</i>	sein, ovaires	Ballestar 2003
<i>CDH1</i>	sein, estomac, leucémie	Bakker 2002; Katoh 2005
<i>GSTP1</i>	sein, prostate, foie	Lin 2001; Song 2002; Pandey 2010; Bakker 2002
<i>MGMT</i>	colon, cerveau, poumon, lymphome	Ballestar 2003; Esteller 1999; Weller 2010
<i>P14(ARF)</i>	colon	Madginier 2001; Martin 2008
<i>P16(INK4A)</i>	colon	Madginier 2001
<i>PRLR</i>	sein	Ballestar 2003
<i>PTPN4</i>	sein	Ballestar 2003
<i>TIMP3</i>	rein, colon	Masson 2010; Rhee 2002
<i>Nrf2</i>	prostate	Yu 2010

Tableau 1 : Gènes hyperméthylés dans les cancers et ciblés par la protéine MBD2

concernant l'expression de *MBD2* intrinsèque aux échantillons issus de tumeurs comparés aux tissus sains.

De plus, les taux d'expression de *MBD2* ne sont corrélés à aucun type histologique ou stade tumoral particuliers (Oue, Kuraoka et al. 2001; Muller, Fiegl et al. 2003; Tryndyak, Kovalchuk et al. 2006).

Ainsi, la protéine MBD2 présente des profils altérés dans les cancers mais aucune variation ne se dégage systématiquement. De plus, il existe aussi des variations au sein du même type de cancer rendant peu probable d'associer une véritable tendance pour un cancer.

Enfin, il n'apparaît pas de corrélation systématique entre le taux d'ARNm *MBD2* et celui des *DNMT* qui se trouve également altéré (Oue, Kuraoka et al. 2001; Sato, Horio et al. 2002; Fang, Cheng et al. 2004; Tryndyak, Kovalchuk et al. 2006; Schmidt, Sedivy et al. 2007). Des diminutions d'expression de *MBD2* peuvent par contre être corrélées à celle de méthyltransférases d'histones (Kn, Bassal et al. 2004; Tryndyak, Kovalchuk et al. 2006).

Aussi, les changements de taux protéique MBD2 ne semblent pas dépendre des dérégulations des enzymes catalysant la méthylation de l'ADN. Par contre, les protéines intervenant dans la machinerie de répression transcriptionnelle semblent altérées de la même façon.

II.F.1.b Les gènes suppresseurs de tumeurs, par l'hyperméthylation de leurs promoteurs, deviennent des cibles privilégiées de MBD2

Il a été vu au cours du chapitre précédent que les cellules cancéreuses présentaient un patron de méthylation d'ADN altéré par rapport aux cellules normales, avec une hypométhylation globale du génome associée à une hyperméthylation locale d'ilots CpG notamment présents au niveau de promoteurs de gènes suppresseurs de tumeurs. Ces ilots méthylés devenaient donc des cibles potentielles pour les protéines à MBD.

La fixation des protéines à MBD sur les promoteurs d'un gène candidat ou d'un ensemble de gènes a surtout été déterminée par une approche directe avec des immunoprécipitations de chromatine (ChIP) ou des ChIP hybridées sur puces à ADN (ChIP-on-chip). Certains gènes suppresseurs de tumeurs ont ainsi été caractérisés comme cibles notamment de la protéine MBD2 dans différents types de cancers (**Tableau 1**). Ces liaisons

peuvent être spécifiques ou non de MBD2 et retrouvées dans un ou plusieurs types cellulaires, comme décrit au cours de la partie précédente ou visualisé dans le **tableau 1**.

II.F.2 Rôles de la protéine MBD2 dans différents cancers

Différentes études ont consisté à évaluer les effets de l'inhibition de *Mbd2* ou *MBD2* *in vivo* et *in vitro* dans des modèles cancéreux ou de prédisposition au cancer.

Les souris nullizygotes *Mbd2*^{-/-} ou hétérozygotes *Mbd2*^{+/-} ne présentent pas d'altérations phénotypiques importantes. Cependant, lorsqu'elles sont croisées avec des souris présentant un fond génétique de prédisposition au cancer intestinal (*Apc*^{Min/+}), il apparaît que le nombre et la taille des tumeurs sont réduits chez les souris *Mbd2*^{-/-}/*Apc*^{Min/+} ou *Mbd2*^{+/-}/*Apc*^{Min/+}. Ces souris ont aussi une durée de vie augmentée par rapport aux témoins. Aussi, *Mbd2* favoriserait le développement et la croissance de la tumeur (Sansom, Berger et al. 2003; Berger and Bird 2005), contrairement à ce qui est observé lors d'expériences similaires croisant les souris *Apc*^{Min/+} et *Mbd4*^{-/-} (Wong, Yang et al. 2002). L'étude réalisée par Phesse et coll. a apporté un complément d'informations à ces travaux en comparant par une analyse globale les transcriptomes des souris *Apc*^{-/-} (invalidation conditionnelle) par rapport aux souris *Apc*^{-/-}/*Mbd2*^{-/-} et contrôles (*Apc* non invalidé) (Phesse, Parry et al. 2008). Il apparaît que certains gènes ciblés par la voie de signalisation Wnt présentent une sur-expression dans les souris *Apc*^{-/-} par rapport aux contrôles mais qui est partiellement corrigée dans les souris *Apc*^{-/-}/*Mbd2*^{-/-}. L'explication, apportée par des ChIP, serait que la protéine *Mbd2* via le recrutement du complexe NuRD réprime la transcription de *Lect2*, un inhibiteur de la voie Wnt qui apparaît être impliquée dans l'inhibition de la prolifération et migration cellulaires.

Le promoteur du gène suppresseur de tumeur *GSTP1* (*Glutathione-S transférase P1*) devient hyperméthylé dans divers cancers (Bakker, Lin et al. 2002; Lin and Nelson 2003) notamment de la prostate (Pandey, Shukla et al.; Lin, Tascilar et al. 2001; Song, Stirzaker et al. 2002), et apparaît majoritairement ciblé par la protéine MBD2 (Lin and Nelson 2003; Stirzaker, Song et al. 2004). L'invalidation de l'ARNm *MBD2* par des ARN anti-sens dans des cellules cancéreuses prostatiques à un stade tardif (Shukeir, Pakneshan et al. 2006) paraît diminuer l'expression de gènes impliqués dans l'apparition de métastases et réduire la capacité d'invasion de ces cellules lorsqu'elles sont mises en chambre de Boyden. Lorsque

ces cellules (et leurs contrôles) sont greffées dans des souris males BALB/c *nu/nu*, cela se traduit *in vivo* par une diminution du volume des tumeurs associée à une diminution de l'angiogenèse. Aussi, la protéine MBD2 jouerait un rôle dans l'invasion tumorale au niveau angiogénique et métastatique des cellules malignes.

Ainsi, la protéine MBD2 paraît impliquée dans la tumorigenèse au niveau du développement, de la croissance en favorisant l'angiogenèse et en empêchant la voie Wnt d'être inhibée, ce qui favorise aussi l'apparition de métastases en agissant sur la capacité d'invasion cellulaire. La protéine MBD2 semble donc importante à tous les stades de la progression tumorale.

II.F.3 La protéine MBD2, nouvelle cible pour un traitement anti-cancéreux ?

Au vu des résultats évoqués ci-dessus et au phénotype observé suite à l'inactivation de *Mbd2*, il paraît possible de proposer la protéine MBD2 comme nouvelle cible thérapeutique anti-cancéreuse. Dans ce cadre, différentes études ont eu lieu sur le principe d'inhibition de l'ARNm *MBD2* par des ARN sens ou anti-sens (Slack, Bovenzi et al. 2002; Ivanov, Lamrihi et al. 2003; Campbell, Bovenzi et al. 2004). *In vitro*, l'injection de ces ARN dans des cellules cancéreuses du sein, colon et poumon paraît limiter la formation de colonies en milieu semi-solide, une caractéristique des cellules tumorales qui ont perdu leur propriété de dépendance à l'ancrage. *In vivo*, ces cellules cancéreuses pulmonaires ou colorectales humaines peuvent former des tumeurs dans des souris immunodéficientes. L'injection d'ARN au sein de ces tumeurs ou dans les souris paraît freiner l'augmentation de leur volume et poids par rapport aux tumeurs non traitées voire fait complètement régresser ces tumeurs. A partir de tumeurs pulmonaires xéno greffées, l'équipe a également mis en évidence que si l'injection de bleomycine, induisant des dommages à l'ADN et bloquant le cycle cellulaire après la phase G2, donne des résultats similaires, la combinaison de ces deux traitements a un effet additionnel (Ivanov, Lamrihi et al. 2003).

De façon intéressante, au cours de ces études, il a été montré que ces injections d'ARN anti-MBD2 n'ont pas d'impact sur la croissance des cellules normales *in vitro* et ne présentent pas d'effet de toxicité *in vivo* (Ivanov, Lamrihi et al. 2003; Campbell, Bovenzi et al. 2004) contrairement aux traitements déméthylant actuellement proposés.

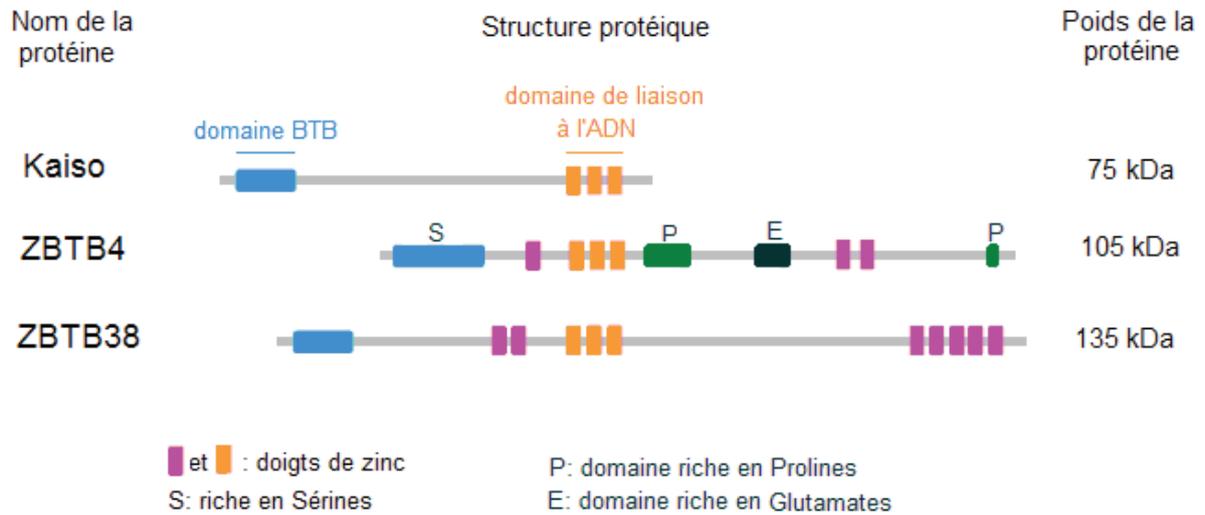


Figure 15 : Schématisation protéique de la famille Kaiso

La famille Kaiso est composée de 3 membres : Kaiso, le membre fondateur, ZBTB4 et ZBTB38. Ces protéines possèdent toutes un domaine de liaison à l'ADN ainsi qu'un domaine BTB (Broad complex, Tramtrak, Bric à brac). Les protéines ZBTB4 et ZBTB38 comprennent d'autres doigts de zinc possèdent des domaines enrichis en certains acides aminés.

Ainsi, la protéine MBD2 paraît être pour le moment, une cible séduisante pour un traitement anti-cancéreux bien plus adapté que celui touchant la méthylation de l'ADN. L'inhibition de *MBD2* pourrait être utilisée en association avec d'autres molécules pour un effet additionnel ou synergique de leurs effets.

III- La famille Kaiso

Chez les mammifères, une seconde famille de protéines capables de lier l'ADN méthylé a été identifiée. Elle est composée de Kaiso, le membre fondateur (Daniel and Reynolds 1999) également appelé ZBTB33, de ZBTB4 et ZBTB38 (Filion, Zhenilo et al. 2006).

III.A Caractérisation des domaines protéiques de la famille Kaiso

Ces protéines appartiennent toutes à la grande famille des facteurs de transcription à doigts de zinc (ZF) et, comme une soixantaine de membres, elles possèdent un domaine BTB/POZ (Broad complex, Tramtrak, Bric à brac/Pox virus and zinc finger) en N-terminal pour une interaction protéine-protéine et un domaine C-terminal constitué de doigts de zinc (**Figure 15**). C'est un motif de 3 doigts de zinc C₂H₂ Kruppel-like en tandem qui caractérise cette famille de protéines (Filion, Zhenilo et al. 2006).

III.A.1 Un domaine en doigts de zinc et jusqu'à deux fonctions

III.A.1.a Un domaine en doigts de zinc conférant une double spécificité de liaison à l'ADN

III.A.1.a.i Identification du domaine de liaison à l'ADN

Kaiso a été identifié suite à la recherche de partenaires de p120-caténine appelée p120^{ctn} (Daniel and Reynolds 1999). Sa séquence protéique contenant 3 doigts de zinc C₂H₂ Kruppel-like adjacents, il a suggéré puis démontré que cette protéine se lie à l'ADN via ce

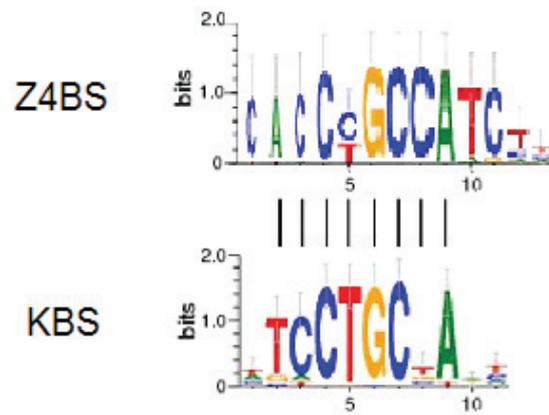


Figure 16 : Motifs des séquences consensus pour la liaison de Kaiso (KBS) et de ZBTB4 (Z4BS) (Figure adaptée de Sasai, Nakao et al., 2010)

motif (Daniel, Spring et al. 2002). Cette protéine pouvant se fixer sur des sites méthylés, d'autres protéines possédant cette propriété ont été recherchés par homologies de séquences de domaines semblables à ce motif. Cela a conduit à la mise en évidence des autres membres de la famille (Filion, Zhenilo et al. 2006), d'une part ZBTB4 et d'autre part ZBTB38 dont les orthologues ont été précédemment décrits chez le rat sous le nom de Zenon (Kiefer, Chatail-Hermitte et al. 2005) et chez la souris sous celui de CIBZ (Sasai, Matsuda et al. 2005).

III.A.1.a.ii Une liaison à l'ADN bimodale

Les 3 doigts de zinc C_2H_2 Kruppel-like adjacents permettent aux protéines de cette famille de lier l'ADN méthylé *in vivo* et *in vitro*, suite à la reconnaissance de 2 sites méthylés consécutifs (Daniel, Spring et al. 2002) pour Kaiso et d'un seul pour ZBTB4 et ZBTB38 (Filion, Zhenilo et al. 2006). Cependant, une autre étude récente (Sasai, Nakao et al.) montre par retard sur gel qu'un contexte séquence spécifique CMGCCATC (M = cytosine méthylée), augmenterait l'affinité de liaison de ces protéines à l'ADN méthylé, tout comme MeCP2 dans une séquence riche en A/T (Klose, Sarraf et al. 2005). A ce moment là, un seul site méthylé symétrique suffirait à la liaison de Kaiso et ZBTB38 tandis que ZBTB4 pourrait fixer un site hémi-méthylé. D'autre part, ce même type d'expérience réalisé à partir de mutants de protéines Kaiso fusionnées avec la GST et incubés avec une séquence cible de la matrilysine, laisse supposer que seuls les doigts de zinc 2 et 3 sont nécessaires pour une liaison de Kaiso à l'ADN méthylé (Daniel, Spring et al. 2002).

De plus, ce même domaine en doigts de zinc permet la liaison des protéines de la famille à des séquences consensus non méthylées. Elles ont été identifiées chez l'Homme pour Kaiso et ZBTB4, chez le rat pour ZBTB38. Il s'agit des séquences CTGCNA (N= un nucléotide) pour Kaiso (Daniel, Spring et al. 2002), C(CouT)GCCATC pour ZBTB4 (Sasai, Nakao et al.) qui peut donc lier le même site que Kaiso et CACCTG pour ZBTB38 (Kiefer, Chatail-Hermitte et al. 2005). Ces séquences spécifiques sont appelées respectivement KBS (Kaiso Binding Site), Z4BS (ZBTB4 Binding Site) et E-box. Les motifs de séquence de Kaiso et ZBTB4 sont très semblables (**Figure 16**) et sont méthylables, ce qui favorise d'autant plus la liaison de ces protéines à ces loci.

On peut donc parler de liaison à l'ADN bimodale pour toutes les protéines de la famille, avec une liaison préférentielle de ZBTB4 aux sites méthylés (Sasai, Nakao et al.) et de Kaiso aux sites non méthylés (Daniel, Spring et al. 2002).

III.A.1.b Un domaine en doigts de zinc pour une interaction protéine-protéine concernant ZBTB4 et ZBTB38

Le rôle principal des 3 doigts de zinc en tandem est de permettre une liaison à l'ADN. Néanmoins, des homo- et hétéro-dimérisations de ZBTB4 et ZBTB38 par ce domaine (Filion, Zhenilo et al. 2006) sont possibles. Si les conséquences ne sont encore pas connues, il est envisageable que cette compétition entre protéines et ADN confère des rôles différents à ces protéines.

Quant à Kaiso, elle n'est pas capable de se lier aux autres membres de la famille et son interaction avec d'autres protéines ne concerne pas ce domaine.

III.A.2 Les autres domaines protéiques

III.A.2.a Les autres doigts de zinc : quelles fonctions ?

Les protéines ZBTB4 et ZBTB38 possèdent d'autres doigts de zinc dans leur domaine C-terminal mais ils ne semblent pas participer pas à la liaison à l'ADN (Filion, Zhenilo et al. 2006). Si leurs rôles sont encore à déterminer, il ne peut pas être exclu que ces domaines puissent participer à des interactions avec d'autres protéines possédant un domaine BTB/POZ. En effet, il a été décrit pour BCL6 et PLZF, deux oncogènes appartenant à la famille BTB/POZ-ZF que le domaine BTB/POZ de chaque protéine interagit avec le domaine ZF de l'autre (Dhordain, Albagli et al. 2000).

III.A.2.b Le domaine BTB/POZ

Toutes les protéines de la famille Kaiso possèdent un domaine BTB/POZ. Ce domaine peut permettre de recruter des complexes de modification de la chromatine pour induire une

répression transcriptionnelle, comme cela a été établi pour d'autres membres de la famille des facteurs de transcription. Cette propriété est retrouvée *in vivo* pour Kaiso chez l'Homme (Prokhortchouk, Hendrich et al. 2001; Yoon, Chan et al. 2003). Ainsi, ce domaine pourrait remplir une fonction équivalente à celle du domaine de répression transcriptionnelle (TRD) des protéines à MBD.

L'implication de ce domaine dans la répression transcriptionnelle a été mise en évidence *in vitro* pour CIBZ/ZBTB38 (Sasai, Matsuda et al. 2005; Filion, Zhenilo et al. 2006) mais pas pour ZBTB4. Il n'est pas à exclure que l'enrichissement en sérines du domaine BTB/POZ de ZBTB4 puisse lui conférer des propriétés particulières par rapport aux autres membres de la famille Kaiso.

Le domaine BTB/POZ est également impliqué dans l'homodimérisation de Kaiso (Daniel and Reynolds 1999), comme de nombreuses autres protéines BTB/POZ – ZF. Si le rôle exact de cette homodimérisation n'est pas connu, il est envisageable qu'elle favorise le recrutement de complexes multiprotéiques plus grands au niveau de l'ADN, ce que ne peut pas permettre l'homodimérisation des autres protéines de la famille.

D'autre part, ce domaine est responsable de l'hétérodimérisation des protéines Zbtb4 et Miz1 (Weber, Marquardt et al. 2008). Il apparaît que Kaiso peut aussi s'hétérodimériser avec des facteurs de transcription à doigts de zinc comme CTCF via ce domaine (Defossez, Kelly et al. 2005) et à d'autres protéines BTB/POZ- ZF comme Znf131 (Donaldson, Nordgaard et al.). Cette découverte concorde avec l'idée que l'hétérodimérisation est un phénomène commun aux protéines POZ même si elles n'interagissent pas toutes entre elles.

III.B Localisations cellulaires et tissulaires des membres de la famille Kaiso

III.B.1 Localisation au sein des compartiments cellulaires

Les protéines ZBTB4 et ZBTB38 présentent une localisation exclusivement nucléaire. En effet, ces protéines, exprimées de façon exogène pour les marquer et les visualiser par immunofluorescence, sont retrouvées au niveau des chromocentres, régions riches en CpG méthylés (Sasai, Nakao et al.; Sasai, Matsuda et al. 2005; Filion, Zhenilo et al. 2006). Ces endroits sont aussi fixés par les protéines à MBD. Dans des cellules humaines, la transfection

transitoire d'ADN codant ces protéines montre que ZBTB4 et ZBTB38 co-localisent partiellement (Filion, Zhenilo et al. 2006), pouvant ainsi leur permettre de s'hétérodimériser. Si aucun domaine de localisation nucléaire (NLS) n'a été identifié pour ces protéines, il apparaît que le glutamate en position 350 est indispensable à la localisation de ZBTB4 au niveau de l'hétérochromatine *in vivo* (Sasai, Nakao et al.).

Par contre, Kaiso n'est pas localisée avec les autres membres de la famille et a une distribution variable entre le noyau et le cytoplasme selon le type cellulaire (Daniel and Reynolds 1999; Kelly, Otchere et al. 2004) et le stade de progression tumorale (Dai, Wang et al.). En effet, plus la cancérogenèse est avancée plus Kaiso aurait tendance à devenir cytoplasmique. Cette protéine est capable de s'importer dans le noyau puisqu'elle a un domaine NLS par lequel elle interagit avec le récepteur d'import nucléaire importine $\alpha 2$ (Kelly, Otchere et al. 2004). Sa répartition complexe dépendrait néanmoins de facteurs du microenvironnement (Soubry, van Hengel et al. 2005). En effet, après injections de cellules tumorales dans des souris immunodéficientes, Kaiso passe du noyau au cytoplasme et la culture *in vitro* de ces xénogreffes montre que Kaiso redevient nucléaire. De plus, il n'est pas clairement établi si son partenaire p120^{ctn}, qui possède un domaine NLS et probablement 2 NES (nuclear export signal) (Dai, Wang et al.; Kelly, Spring et al. 2004), joue un rôle clé dans sa répartition intracellulaire. Cela permet également de suggérer que Kaiso puisse être séquestrée par des protéines cytosoliques et/ou exportée par des protéines nucléaires.

III.B.2 Localisation au sein des lignées cellulaires et des tissus

Au vu des résultats obtenus chez les rongeurs, les 3 protéines de la famille paraissent avoir une expression ubiquitaire avec cependant quelques différences entre les tissus. En effet, l'orthologue Zenon de ZBTB38 est présent essentiellement au niveau du cerveau et des tissus neuroendocriniens (Kiefer, Chatail-Hermitte et al. 2005), de même pour CIBZ également détectée à de plus faibles taux dans tous les autres tissus adultes (Sasai, Matsuda et al. 2005), suggérant une prédominance de ZBTB38 dans le rein et les neurones. Les protéines Kaiso et ZBTB4 semblent présenter des patrons d'expressions similaires (Daniel and Reynolds 1999; Filion, Zhenilo et al. 2006) avec une prépondérance notamment dans le cerveau (Della Ragione, Tiunova et al. 2006; Filion, Zhenilo et al. 2006), contrairement à ce qui avait été

initialement décrit pour Kaiso (Daniel and Reynolds 1999). Il est à noter que ZBTB4 n'est pas détectée dans les embryons.

Chez l'Homme, les études d'expression ont été principalement effectuées dans des lignées cellulaires. Ainsi, Kaiso est exprimée dans des lignées cellulaires cancéreuses de sein, rein et col de l'utérus. Elle est retrouvée dans des lignées et tissus de poumon (Dai, Wang et al.) et de colon (Lopes, Valls et al. 2008), ainsi que dans le cerveau (Della Ragione, Tiunova et al. 2006). La protéine ZBTB4 est exprimée dans les cellules HeLa comme ZBTB38, paraît présenter un taux diminué par rapport au tissu normal dans les stades avancés de neuroblastomes ainsi que dans de multiples tumeurs solides comme celles touchant le sein, la prostate (Weber, Marquardt et al. 2008).

III.C Rôles des protéines de la famille Kaiso dans la régulation génique

III.C.1 Une dualité fonctionnelle pour Kaiso et ZBTB38

III.C.1.a Kaiso et ZBTB38, des répresseurs transcriptionnels

III.C.1.a.i Le répresseur Kaiso se lie sur des séquences nucléotidiques via son domaine à doigts de zinc

- ***Indépendamment de la méthylation de l'ADN***

Kaiso, en se liant sur les sites KBS non méthylés, peut initier une répression transcriptionnelle de deux manières.

Cette fixation permettrait le recrutement de partenaires répresseurs encore non connus et serait modulée par d'autres facteurs protéiques, avec pour conséquence une régulation transcriptionnelle complexe. En effet, le co-facteur p120^{ctn} dont le site de fixation à Kaiso englobe les 3 doigts de zinc, peut empêcher la liaison de Kaiso à l'ADN avec pour conséquence une diminution de la répression du gène *Wnt11* chez le Xénope, la souris et l'Homme (Kim, Park et al. 2004). Il peut aussi agir en partenariat avec la β -caténine pour bloquer totalement l'action de Kaiso au niveau du gène *matrilysine* codant la protéine MMP-7 (métalloprotéinase de matrice 7) (**Figure 17A**) (Spring, Kelly et al. 2005). L'activité de p120^{ctn} peut elle-même être influencée par d'autres facteurs appartenant à la voie de

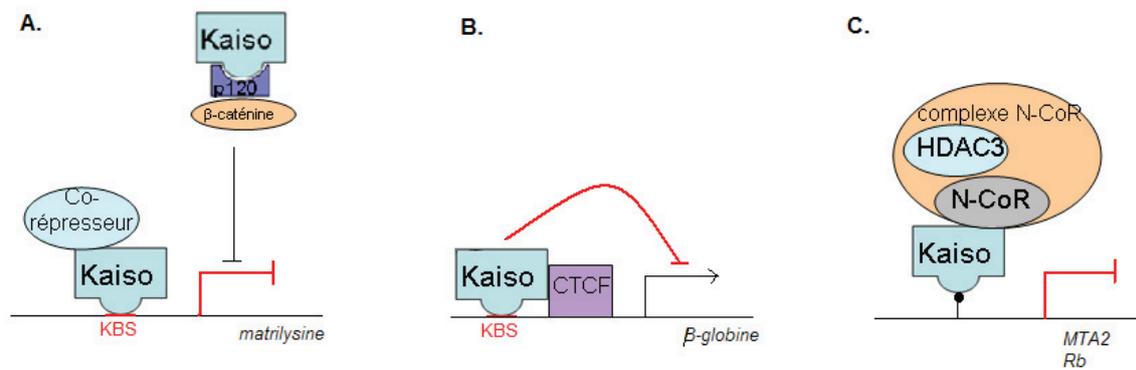


Figure 17 : Répressions transcriptionnelles dépendantes de la méthylation de l'ADN initiées par la protéine Kaiso

signalisation Wnt tels que Frodo lui-même régulé par des protéines Dishevelled (Na, Lykke-Andersen et al. 2007).

A l'inverse, la liaison de Kaiso à des sites consensus non méthylés pourrait inhiber l'activation transcriptionnelle de gènes et donc induire une répression de façon indirecte. Ainsi, en 5' de l'insulateur de la β -globine humaine, Kaiso peut se fixer au site KBS situé à proximité de celui de l'insulateur CTCF et interagir avec lui via son domaine BTB/POZ et ainsi empêcher la transcription (Defossez, Kelly et al. 2005) (**Figure 17B**).

- *Après reconnaissance de CpG méthylés*

Des expériences de retard sur gel et de spectrométrie de masse ont permis d'identifier un complexe protéique contenant Kaiso (Prokhortchouk, Hendrich et al. 2001) et lié à une partie intronique méthylée du gène *S100A4*. Un système de gène rapporteur ainsi que des immunoprécipitations de chromatine ont permis de montrer la fixation de ce complexe via Kaiso par son domaine en doigts de zinc à l'ADN méthylé ainsi que la nécessité de son domaine POZ pour réprimer le gène. Ce complexe répresseur ne contient pas d'HDAC mais certaines protéines appartenant au complexe MeCP1. Enfin, il apparaît que ce complexe est absent dans les cellules HeLa.

Aussi, des expériences similaires ont été menées dans ces cellules afin d'identifier d'autres complexes répresseurs. Kaiso apparaît alors faire partie d'un complexe N-CoR (Yoon, Chan et al. 2003) recruté au niveau du gène *MTA2* via la liaison de Kaiso à l'ADN méthylé. De plus, cette protéine recrute le reste du complexe par son domaine POZ, comme le montre des expériences de chromatographie d'affinité « GST pull-down » ce qui permet une répression, corrélée à une désacétylation par HDAC3 et une méthylation de la lysine 9 de l'histone H3 (**Figure 17C**).

Ce même complexe répresseur est retrouvé pour l'inhibition de la transcription du gène retinoblastome (Rb) lorsque son promoteur devient méthylé au cours de la tumorigenèse. Kaiso remplace alors l'insulateur CTCF, qui permettait une activation transcriptionnelle, sur son site contenant le motif CGCG (De La Rosa-Velazquez, Rincon-Arano et al. 2007). D'autres cibles subissant une hyperméthylation aberrante dans les cancers ont été identifiées (Lopes, Valls et al. 2008) comme les gènes *CDKN2A*, *HIC-1* et *MGMT* pour lesquels les mécanismes de répression ne sont pas connus.

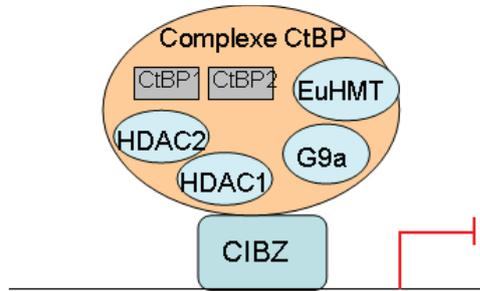


Figure 18 : Seul cas de répression transcriptionnelle rapporté pour CIBZ

Ainsi, après fixation sur une séquence méthylée, Kaiso semble capable d'initier une répression transcriptionnelle en recrutant des complexes répresseurs contenant ou non des désacétylases d'histones.

III.C.1.a.ii Le répresseur Kaiso se fixe à des protéines via son domaine BTB/POZ

L'interaction de Kaiso par son domaine BTB/POZ avec d'autres protéines impliquées directement dans l'activation de l'expression génique peut inhiber leur action. En effet, par ce domaine, Kaiso peut s'hétérodimériser avec Znf131, un membre de la famille BTB/POZ-ZF identifié comme activateur transcriptionnel (Donaldson, Nordgaard et al.). Un système de gène rapporteur sous l'influence d'un promoteur contenant 4 sites de fixation de Znf131 a permis de révéler que cette association empêche Znf131 d'agir, par un mécanisme qu'il reste encore à définir. Il est à noter que l'action de Kaiso peut être diminuée par l'introduction de la protéine p120^{cm}, rendant la régulation génique encore plus complexe.

Ainsi, il apparaît que Kaiso peut intervenir sur l'expression génique indépendamment de sa fixation à une séquence nucléotidique.

III.C.1.a.iii CIBZ/ZBTB38 et le co-répresseur CtBP

Il a été suggéré par les expériences de Filion et de ses collaborateurs que ZBTB38 agisse comme un répresseur transcriptionnel (Filion, Zhenilo et al. 2006) mais aucune démonstration n'en a été faite. Par contre, un cas d'inhibition de la transcription initiée par l'orthologue CIBZ et impliquant deux domaines a été rapporté *in vitro*. Ainsi, par un système de gène rapporteur luciférase et un traitement à la TSA inhibant l'activité des HDAC, il a été mis en évidence que CIBZ via ses domaines BTB/POZ et RD2 (repression domain 2) induit une répression transcriptionnelle dépendante des HDAC. Des co-immunoprécipitations ont montré qu'en fait cette inhibition est permise par l'interaction entre CIBZ et CtBP (Sasai, Matsuda et al. 2005), un complexe répresseur constitué notamment de désacétylases (HDAC1 et 2) et méthyltransférases d'histones (G9a et EuHMT) (Shi, Sawada et al. 2003) (**Figure 18**). Le complexe CIBZ/CtBP est retrouvé au niveau des foyers péricentromériques, suggérant

qu'il lie l'hétérochromatine. Il reste à démontrer si cette fixation est dépendante ou non de la méthylation de l'ADN.

III.C.1.b Kaiso et ZBTB38, également des activateurs transcriptionnels

Si Kaiso semble jouer un rôle important dans la répression transcriptionnelle, un cas d'activation a été rapporté. Ainsi, il a été montré par co-précipitation que Kaiso lie la δ -caténine et par immunoprécipitation de chromatine que ce complexe fixe par Kaiso le site KBS présent sur le promoteur du gène de la rapsyne dans des myotubes de souris et poulet (Rodova, Kelly et al. 2004). Le partenaire de Kaiso semble donc jouer un rôle déterminant puisque cette association permet une activation de la transcription dans le muscle.

ZBTB38 semble jouer un rôle différent de Kaiso. En effet, aucun cas de répression transcriptionnelle n'a été rapporté suite à la fixation de ZBTB38 au motif E-box (séquence consensus non méthylée). Par contre, un cas d'activation transcriptionnelle induite par la liaison de Zenon sur le promoteur du gène tyrosine hydroxylase (TH) chez le rat (Kiefer, Chatail-Hermitte et al. 2005) et n'impliquant pas le domaine BTB/POZ a été décrit. Or, chez l'Homme, une répression transcriptionnelle de ce gène a été mise en évidence suite à la liaison de Kaiso sur le premier exon méthylé (Aranyi, Faucheux et al. 2005). Aussi, Kaiso et ZBTB38 pourraient partager des cibles mais la liaison de l'une ou l'autre dépendrait de l'état de méthylation de l'ADN, du type cellulaire et/ou de l'espèce considérée et aboutirait à des effets opposés sur la transcription.

D'autre part, suite aux expériences de gène rapporteur luciférase sous contrôle d'un promoteur contenant le domaine de liaison GAL4 fusionné à différentes formes tronquées de CIBZ, il a été émis l'hypothèse qu'un domaine d'activation existe au sein de cette protéine (Sasai, Matsuda et al. 2005).

III.C.2 ZBTB4 et son activité répressive

De même que ZBTB38, il apparaît *in vitro* que ZBTB4 intervient dans la répression transcriptionnelle (Filion, Zhenilo et al. 2006). Un cas a été décrit *in vivo* concernant le gène *P21CIP1* dont l'expression est initialement induite par la fixation de Miz1, une protéine à

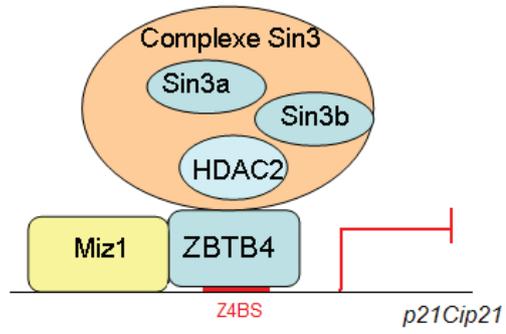


Figure 19 : ZBTB4 est un répresseur transcriptionnel de *p21Cip21*

domaine BTB/POZ. En effet, l'hétérodimérisation via les domaines BTB/POZ de Zbtb4 et Miz1 ainsi que la fixation de Zbtb4 à des sites similaires à KBS répriment la transcription de ce gène lorsque p53 est activé (Weber, Marquardt et al. 2008). Zbtb4 a donc un rôle de répresseur transcriptionnel. Des expériences de co-précipitations montrent qu'elle remplit ce rôle en recrutant comme partenaire un complexe Sin3 contenant Sin3a, Sin3b et au moins HDAC2 (**Figure 19**).

III.D La famille Kaiso et son implication dans les grandes fonctions biologiques chez les vertébrés

III.D.1 Kaiso, ZBTB38 et le développement embryonnaire

Chez l'embryon de Xénope, l'inhibition de la traduction protéique de Kaiso par injections d'oligonucléotides morpholino à un stade précoce conduit à un retard lors de la blastulation puis à une mort par apoptose au stade neurula, conséquences de la réactivation transcriptionnelle de gènes méthylés impliqués dans la voie du développement (Ruzov, Dunican et al. 2004; Ruzov, Savitskaya et al. 2009). En revanche, chez la souris, l'inactivation du gène codant Kaiso, par un système flox-Cre, n'entraîne ni létalité ni altération phénotypique observable (Prokhortchouk, Sansom et al. 2006). Kaiso ne semble donc pas indispensable au cours du développement embryonnaire chez les vertébrés. Par contre, Kaiso et p120^{ctn} paraissent impliqués dans la régulation du gène Wnt11 participant au contrôle de la morphogénèse chez les vertébrés (Kim, Park et al. 2004).

CIBZ peut recruter comme co-répresseur le complexe CtBP via sa fixation à CtBP1 et CtBP2. Or les invalidations des gènes codant pour ces protéines chez la souris montrent qu'elles jouent un rôle fondamental dans le développement chez les mammifères (Hildebrand and Soriano 2002). Aussi, il n'est pas exclu que ZBTB38 joue un rôle dans ce processus biologique.

III.D.2 ZBTB4, ZBTB38 et l'apoptose

L'orthologue murin CIBZ de ZBTB38 possède deux sites de reconnaissance de la caspase-3 qui lui permettent d'interagir avec cette protéine (Oikawa, Matsuda et al. 2008). De plus, l'inactivation par ARN interférence de l'ARNm de *CIBZ* dans les cellules non cancéreuses C2C12 conduit d'une part à l'activation de la caspase-3 et au clivage de la protéine PARP, deux marqueurs caractéristiques de l'apoptose et d'autre part à l'augmentation de l'expression des protéines p21 et p27 associées à la voie apoptotique dépendante de p53. Il reste à établir si ces altérations d'expression sont le reflet d'une interaction directe ou indirecte entre CIBZ et les séquences régulatrices des gènes codant pour ces protéines. Il est également à noter que CIBZ interagit avec le co-répresseur CtBP, un complexe impliqué dans la répression de facteurs pro-apoptotiques par une voie ne faisant pas intervenir p53. Aussi, par ces deux mécanismes, CIBZ paraît jouer un rôle essentiel dans la protection cellulaire vis-à-vis de l'apoptose.

Il a également été démontré que *Zbtb4* se lie au promoteur non méthylé du gène *p21* (Weber, Marquardt et al. 2008). Aussi, il serait intéressant de voir si cette cible est commune à tous les membres de la famille Kaiso, d'autant plus que *Zbtb4* lie le site reconnu par Kaiso.

III.E Kaiso et la cancérogenèse

La majorité des protéines possédant un domaine POZ sont des répresseurs transcriptionnels identifiés par rapport à leur implication dans les cancers, comme BCL6 dans le lymphome non-hodgkinien (Lo Coco, Ye et al. 1994). De plus, la localisation de Kaiso dépendrait de son microenvironnement tumoral (Soubry, van Hengel et al. 2005). Aussi, la protéine Kaiso a été étudiée dans différents modèles murins présentant une prédisposition tumorale ainsi que dans différents cancers humains.

Des souris possédant une inactivation du gène codant pour Kaiso (Prokhortchouk, Sansom et al. 2006) ont été croisées avec le modèle murin *Apc*^{Min/+} présentant une prédisposition aux tumeurs intestinales. Il apparaît que la durée de vie est augmentée et la taille des polypes diminuée par rapport aux souris *Apc*^{Min/+}, montrant que Kaiso est impliquée dans la tumorigénèse intestinale. De plus, la similitude des indices apoptotiques et mitotiques entre les tumeurs issues de souris *Apc*^{Min/+} et *Kaiso*^{-/-}/*Apc*^{Min/+} suggère que Kaiso

n'interviendrait pas au cours de la croissance tumorale mais à des stades plus précoces. De façon intéressante, il a été montré que Kaiso régule différentes protéines impliquées dans la voie de signalisation Wnt. Or c'est une voie importante pour la programmation de la différenciation de l'épithélium intestinal. Par contre, il est à noter que l'absence de Kaiso dans les souris *Apc*^{Min/+} aboutit à des effets plus modérés que l'absence de *Mbd2* ou *Dnmt1*.

De plus, dans le modèle murin *Muc2*^{-/-} prédisposant aux tumeurs colorectales, une augmentation de l'expression de Kaiso est observée dans les tumeurs par rapport à la muqueuse normale. Chez l'Homme, cette même comparaison montre que Kaiso est retrouvée dans tous les échantillons à des taux variables, ne permettant d'établir aucune différence entre la tumeur et la muqueuse normale (Prokhortchouk, Sansom et al. 2006). Une autre étude portant sur les tumeurs humaines du colon (Lopes, Valls et al. 2008) a permis d'identifier comme nouvelles cibles de Kaiso les gènes *CDKN2A*, *HIC-1* et *MGMT* hyperméthylés lors de la tumorigenèse. L'inactivation de l'ARNm de Kaiso par ARN interférence dans les cellules tumorales du colon favorise un arrêt du cycle cellulaire et une mort cellulaire, supposant ainsi que Kaiso facilite notamment la prolifération des cellules cancéreuses.

D'autre part, Kaiso cible le gène *matrilysine* codant la protéine MMP7 et fait intervenir comme partenaire p120^{ctn} pour la répression transcriptionnelle. Or, ce gène est décrit pour être une cible du complexe β -caténine/TCF appartenant à la voie de signalisation canonique Wnt (Spring, Kelly et al. 2005). De plus, l'augmentation du taux protéique de cette métalloprotéinase de la matrice est directement associée aux tumeurs métastatiques dans les cancers du colon et du sein (Rudolph-Owen, Chan et al. 1998). Aussi, il semble exister un lien entre Kaiso et voie Wnt et donc probablement entre Kaiso et la tumorigenèse impliquant la voie Wnt.

III.F Redondance fonctionnelle entre protéines liant l'ADN méthylé ?

III.F.1 Redondance entre les différents membres de la famille Kaiso ?

Il a été vu précédemment que l'inactivation de *Kaiso* chez la souris ne conduit ni à une létalité ni à un phénotype altéré contrairement à ce qui est observé chez le Xénope. Ceci laisse donc supposer qu'il existe une redondance fonctionnelle chez les mammifères entre Kaiso et

d'autres protéines ciblant l'ADN méthylé. Les invalidations des gènes codant pour les protéines ZBTB4 et ZBTB38 ainsi que la création de double ou triple mutants pourraient permettre de répondre à cette question pour la famille Kaiso.

De plus, ce sont des protéines ubiquitaires ce qui est en faveur d'un partage de cibles comme p21 éventuellement. ZBTB4 et ZBTB38 peuvent co-localiser ce qui faciliterait leur remplacement mutuel ou leur entrée en compétition pour un site donné. Si elles remplissaient des rôles communs, il serait même envisageable que dans certains cas leur hétérodimérisation soit équivalente à leurs homodimérisations.

Dans le cas de séquences non méthylées, le site de liaison à l'ADN KBS représente un cas particulier de Z4BS. Ceci suggère donc que, *in vivo*, ZBTB4 peut se distribuer comme Kaiso et en plus se fixer sur des sites spécifiques. D'autre part, le site E-box est très différent de KBS et Z4BS, ce qui conférerait aussi une spécificité de liaison à ZBTB38.

Enfin, Kaiso se distingue des autres membres de la famille de part sa localisation qui peut être cytoplasmique ainsi que par le fait qu'elle ne se dimérise pas avec eux. Ceci suggère donc un ou des rôle(s) spécifique(s) pour cette protéine.

Aussi, s'il existait une redondance fonctionnelle entre les membres de la famille Kaiso, elle serait associée à des spécificités de cibles, comme pour les protéines à MBD.

III.F.2 Redondance entre les différentes familles de protéines liant l'ADN méthylé ?

III.F.2.a Quelques pistes en faveur de cette hypothèse

Les protéines à MBD, à l'exception de MBD3, et les membres de la famille Kaiso partagent toutes la propriété de lier l'ADN méthylé. De plus, il a été vu précédemment que les invalidations de gènes codant les protéines à MBD, à l'exception de MBD3, (Guy, Hendrich et al. 2001; Hendrich, Guy et al. 2001; Zhao, Ueba et al. 2003) et Kaiso (Prokhortchouk, Sansom et al. 2006) n'entraînent pas de phénotype létal. Il a donc été proposé que toutes ces protéines partagent des fonctions communes.

Aussi des triples mutants *Mecp2*^{-/-}, *Mbd2*^{-/-} et *Zbtb33*^{-/-} ont été créés (Martin Caballero, Hansen et al. 2009) et comparés aux simples et doubles mutants. Il apparaît que tous les mutants sont viables et ne présentent pas d'aggravation phénotypique sauf pour les mutants *Mecp2*^{-/-}/*Mbd2*^{-/-} et *Mecp2*^{-/-}/*Mbd2*^{-/-}/*Zbtb33*^{-/-} par rapport mutants *Mecp2*^{-/-} avec une moyenne de vie diminuée. De plus, aucune des protéines des deux familles n'est surexprimée dans les lignées cellulaires souches neurales extraites des simples, doubles et triples mutants, suggérant qu'il n'a pas de compensation en terme de concentration. Par contre, au cours de la différenciation neurale, un retard puis une compensation sont observés dans les triples mutants par rapport aux cellules non mutées ou simples mutants. Ainsi, cette étude est la première à suggérer, *in vitro*, une certaine redondance fonctionnelle des protéines liant l'ADN méthylé dans des cellules de mammifères.

De plus, des parallèles peuvent être faits concernant la fixation de toutes ces protéines à l'ADN. En effet, MBD4 et ZBTB4 sont capables de reconnaître l'ADN hémi-méthylé et les membres de la famille Kaiso peuvent se lier à l'ADN méthylé ou non méthylé, tout comme MBD1. Ces homologies de liaisons pourraient donc être en faveur d'une redondance fonctionnelle.

Enfin, ces protéines liant l'ADN méthylé semblent capables de lier les mêmes sites. En effet, il a été vu par CHIP que MeCP2 ou le complexe MeCP1 peuvent occuper le même site sur le promoteur de *SYBL1*, au même titre que Kaiso (Matarazzo, De Bonis et al. 2007).

III.F.2.b MeCP2, Kaiso et le syndrome de Rett

Le syndrome de Rett chez l'Homme s'explique en grande majorité par des mutations du gène MeCP2 (Amir, Van den Veyver et al. 1999). Etant donné que certains patients ne répondent pas à ce critère, il a été émis l'idée que d'autres altérations interviennent au niveau du gène codant pour Kaiso, une autre protéine liant l'ADN méthylé et abondante dans le cerveau. Cette hypothèse n'a pas pu être validée (Della Ragione, Tiunova et al. 2006), excluant pour le moment une redondance fonctionnelle entre Kaiso et MeCP2 dans ce tissu.

III.F.2.c Conclusion : la lecture de la méthylation de l'ADN est un mécanisme complexe

Les différentes familles de protéines liant l'ADN méthylé semblent donc très majoritairement capables de « lire » le signal de méthylation de l'ADN. Mais cette lecture est rendue complexe par l'existence d'un ensemble de mécanismes comme les spécificités de cibles associées à des redondances fonctionnelles entre ces protéines, la présence éventuelle de co-facteurs dirigeant ces protéines et de régulations post-traductionnelles modifiant leur activité ou recrutement.

Après lecture, une « traduction » est faite essentiellement par recrutement de complexes enzymatiques induisant le remodelage de la chromatine. Au cours du chapitre suivant, nous verrons comment la chromatine est organisée et quelles protéines sont responsables de ses compactations et décondensations et plus particulièrement quels changements/acteurs participent à la répression transcriptionnelle.

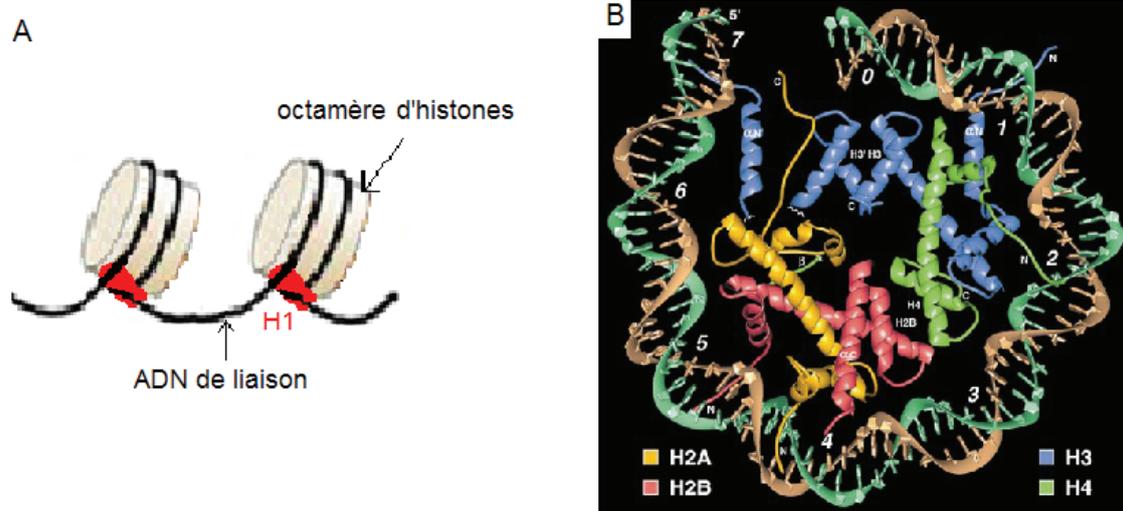


Figure 20 : Formation d'un nucléosome, l'unité fondamentale de la chromatine

- A) Représentation schématique de 2 nucléosomes adjacents. Chaque nucléosome est formé d'une histone H1, d'ADN de liaison et d'un octamère d'histones autour duquel s'enroule 146 paires de bases d'ADN
- B) Structure d'un octamère d'histones, obtenue par cristallographie à une résolution de 2,8 angströms. (Figure issue de Luger, Mader et al., 1997). Un octamère d'histones est formé d'un tétramère central H3-H4 autour duquel se trouvent les dimères H2A-H2B

Une traduction chromatinienne pour une répression transcriptionnelle

I- Les histones, protéines structurales de la chromatine

I.A Le nucléosome : unité fondamentale de la chromatine

L'ADN, de part sa longueur, nécessite d'être compacté au sein du noyau de la cellule. Cette structure condensée, appelée chromatine, a pour unité de base le nucléosome identifié par cristallographie. Chez l'Homme, les 32 millions de nucléosomes (Weidemann, Wachsmuth et al. 2003) sont chacun formés d'un corps protéique autour duquel s'enroule l'ADN en super-hélice (Noll 1974), d'une histone de liaison H1 (Oudet, Gross-Bellard et al. 1975) présente au point d'entrée/sortie des éléments nucléoprotéiques et de l'ADN reliant ces derniers (**Figure 20A**). L'enroulement de l'ADN autour du corps protéique aboutit à une compaction de la chromatine qui peut encore plus se condenser selon l'état de repliement de l'ADN de liaison.

I.B Le corps nucléoprotéique : unité fondamentale du nucléosome

Le corps protéique correspond à un octamère d'histones H2A, H2B, H3 et H4 présentes chacune en 2 copies. L'étude d'histones extraites de chromatine en milieu acide et filtrées sur gel (Kornberg and Thomas 1974) ainsi que la migration sur gel polyacrylamide-SDS de chromatine fixée par un agent bivalent (Thomas and Kornberg 1975) ont permis d'établir que les histones de cet octamère se répartissent en un tétramère central H3-H4 autour duquel se trouvent les dimères H2A-H2B (**Figure 20B**).

Des expériences de diffraction par rayons X de plus en plus résolutive ont par la suite apporté des précisions sur la structure des histones. Chaque histone est constituée d'un domaine globulaire très conservé entre espèces (Luger, Mader et al. 1997) contrairement aux extensions et queues C- et N- terminales (Xie, Kokubo et al. 1996). De plus, il apparaît que

146 paires de bases d'ADN (Arents, Burlingame et al. 1991) s'enroulent en environ 1,7 tours (Luger, Mader et al. 1997; Bednar, Horowitz et al. 1998) autour de l'octamère d'histones.

II- Rôle prépondérant des histones dans la stabilisation de la chromatine

Pour que la chromatine reste stable, il est nécessaire de maintenir la conformation de l'octamère d'histones et de favoriser l'enroulement de l'ADN autour de ce dernier. De plus, les histones de liaison H1, initialement décrites comme répresseurs de l'acétylation (Herrera, West et al. 2000), paraissent aussi contribuer au maintien de la structure chromatinienne.

Ainsi, la composition en acides aminés des histones favorise la présence de charges à leurs surfaces et *a fortiori* leurs interactions (Luger, Mader et al. 1997). De plus, l'étude du nucléosome par diffraction par rayons X a permis de déterminer, en plus de sa structure tridimensionnelle, les multiples liaisons entre l'ADN et les dimères et tétramères d'histones (Wu and Davey; Klug, Rhodes et al. 1980; Luger, Mader et al. 1997; Davey and Richmond 2002). Il n'est pas exclu que la présence de séquences riches en GC (Dickerson, Goodsell et al. 1994; Tillo and Hughes 2009) et en AT (Anderson and Widom 2001) respectivement dans les grand et petit sillons d'ADN favorisent le maintien de l'ADN autour de l'octamère.

Tous les composants du nucléosome, et particulièrement les histones, participent à la stabilité chromatinienne. La chromatine n'est pourtant pas une structure figée. Elle répond à un mécanisme dynamique de compaction/décompaction déterminant l'accessibilité de l'ADN à la machinerie transcriptionnelle et donc sélectionnant les gènes à exprimer et réprimer. Ces changements chromatiniens interviennent aussi dans la réplication, la recombinaison et la réparation de l'ADN. L'enroulement de l'ADN au sein du nucléosome, le degré de repliement de l'ADN de liaison mais aussi les modifications post-traductionnelles des histones participent à ces phénomènes.

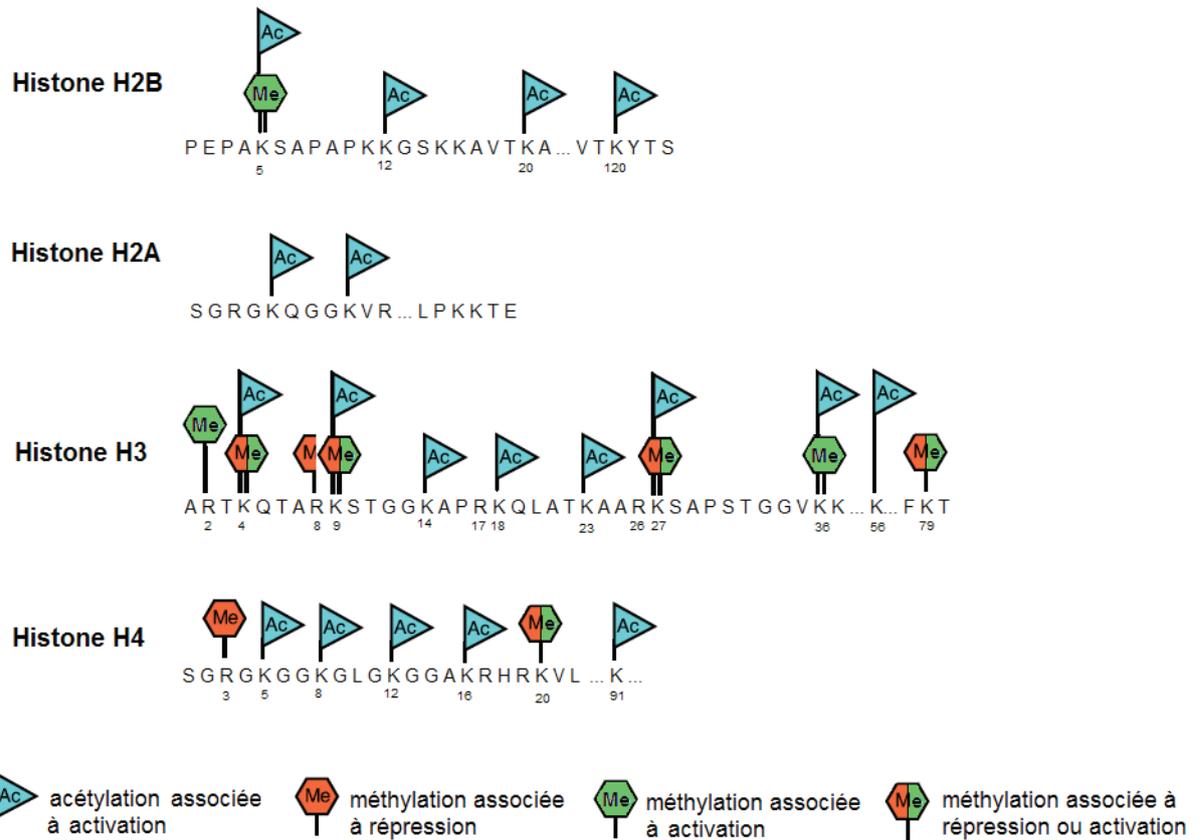


Figure 21 : La méthylation et l'acétylation, deux modifications post-traductionnelles d'histones associées à la régulation transcriptionnelle

Les positions des marques de méthylation et d'acétylation pour les différentes histones et leurs conséquences sur la régulation transcriptionnelle sont répertoriées. Les marques présentes sur un même site sont mutuellement exclusives.

II.A Des modifications post-traductionnelles des histones associées à la régulation génique

Il existe quatre modifications post-traductionnelles majeures des histones c'est-à-dire l'acétylation, la méthylation, la phosphorylation et l'ubiquitinylation auxquelles sont associées la sumoylation, l'ADP-ribosylation, la biotinylation, l'isomérisation de la proline et la citrullination. Le nombre de ces marques ainsi que leurs différentes associations laissent supposer qu'il existe un code des histones (Strahl and Allis 2000) qui, une fois lu, aboutira à la compaction ou à l'ouverture de la chromatine. La complexité de ce code (Berger 2007) serait d'ailleurs renforcée par l'existence de variants d'histones (Franklin and Zweidler 1977) ayant eux-mêmes des isoformes comme pour H2A.Z (Matsuda, Hori et al.).

Toutes les modifications post-traductionnelles des histones sont associées à la régulation transcriptionnelle (Kouzarides 2007) comme la phosphorylation de la sérine S10 de l'histone H3 impliquée dans l'activation transcriptionnelle (Winter, Simboeck et al. 2008). Dans ce chapitre seront abordées l'acétylation/désacétylation et la méthylation/déméthylation des histones. En effet, la désacétylation et la méthylation des histones sont les seules marques associées à la traduction de la méthylation de l'ADN en un état chromatinien approprié pour une répression transcriptionnelle (**Figure 21**).

II.B L'acétylation et la désacétylation des histones

II.B.1 Identification du groupement acétyle sur les lysines des histones

Après avoir identifié des groupements acétyles au niveau d'histones isolées de thymus de veau (Johns, Phillips et al. 1961), il a été démontré que ces groupements se retrouvent préférentiellement dans des fractions histones riches en lysines (K) (Phillips en 1963).

Si toutes les histones portent des marques d'acétylation, elles sont surtout observées pour H3 et H4 (Turner 1993; Felsenfeld and Groudine 2003; Bonenfant, Towbin et al. 2007).

II.B.2 Le jeu des HAT et HDAC

II.B.2.a L'acétylation, une modification catalysée par des HAT

Il existe de nombreuses acétyltransférases réparties en catégories en fonction de leurs localisations cellulaires. Parmi celles-ci, on retrouve 3 familles (GNAT, MYST et p300/CBP) faisant partie de la catégorie A c'est-à-dire se trouvant dans le noyau et responsables de l'acétylation des histones (Sternier and Berger 2000; Roth, Denu et al. 2001).

Chaque acétyltransférase d'histones (HAT) catalyse le transfert d'un groupement acétyle présent dans l'acetyl coenzyme A sur un ou plusieurs résidus lysines. A titre d'exemples, la HAT GCN5, la première HAT de classe A identifiée chez l'Homme (Candau, Moore et al. 1996), intervient sur la lysine 36 de l'histone H3 (H3K36) (Morris, Rao et al. 2007), les sites H3K9, H3K14 et H3K18 (Kouzarides 2007). Cela aboutit à la neutralisation de la charge positive des lysines, ce qui aurait pour conséquences de réduire les contacts avec l'ADN chargé négativement (Hong, Schroth et al. 1993; Puig, Belles et al. 1998), de changer les interactions entre les histones et avec d'autres facteurs impliqués dans la régulation transcriptionnelle. La structure des nucléosomes serait alors altérée, la chromatine plus ouverte et donc plus permissive à la machinerie transcriptionnelle. Ceci est en accord avec les résultats d'un séquençage haut-débit de fragments de chromatine digérée par la nucléase micrococcocale révélant une absence de nucléosomes juste après le site d'initiation de la transcription sur lequel se trouve l'ARN polymérase II (Schones, Cui et al. 2008). De plus, d'autres études montrent en utilisant le modèle levure *Saccharomyces cerevisiae* que cette perte nucléosomique est corrélée à l'activation génique (Boeger, Griesenbeck et al. 2003) et que l'hyperacétylation précède ce remodelage chromatinien (Reinke and Horz 2003).

Il est à noter que les HAT peuvent aussi intervenir dans la régulation transcriptionnelle sans entraîner de modifications de la chromatine. En effet, p300/CBP peut acétyler p53 pour augmenter son affinité de liaison à l'ADN (Gu, Shi et al. 1997) ou agir en tant que pont protéique entre la β -caténine et TCF4.

II.B.2.b Les HDAC catalysent l'excision des groupements acétyles

Chez les mammifères, l'excision des groupements acétyles sur les histones est catalysée par les désacétylases d'histones (HDAC). Cette activité désacétylase a été mise en évidence pour HDAC1, la première HDAC identifiée, à partir de fractions protéiques nucléaires de thymus bovin incubées sur une matrice sur laquelle est fixée un dérivé de la trapoxine, un inhibiteur de cette activité (Taunton, Hassig et al. 1996).

Les HDAC, au nombre de 18 actuellement, sont réparties en 4 classes, en fonction de leur degré d'homologie aux orthologues Rpd3 et Hda1 présents chez la levure *S.cerevisiae*, avec la classe I qui contient les HDAC1, 2, 3 et 8 (Yang and Seto 2008). Les HDAC1, 2 et 3 présentent une forte homologie. Elles possèdent un domaine catalytique pour l'excision du groupement acétyle, un domaine C-terminal avec des sites phosphorylables CK2 en tandem pour HDAC1 et 2 alors que HDAC3 n'a qu'un site CK2. Le domaine C-terminal de HDAC1 et 2 paraît même leur conférer une spécificité de cible (Hong, Derfoul et al. 2009).

Comme nous l'avons vu précédemment au cours du chapitre 2, pour induire une répression transcriptionnelle, après recrutement par les protéines à MBD, les HDAC1 et 2 agissent au sein de complexes ATP-dépendant remodelant la chromatine (Yang and Seto 2008), ce qui pourrait être aussi le cas pour HDAC3. Les conséquences de leurs actions sur la régulation de la transcription des gènes doivent être à l'inverse de celles observées après ajout du groupement acétyle par les HAT. Leur activité enzymatique pourrait être régulée par des modifications post-traductionnelles. L'activité enzymatique de HDAC3 est empêchée par la déphosphorylation de la sérine 424 (sur le site CK2) par PP4_c (Zhang, Ozawa et al. 2005) alors que celle de HDAC1 et sans doute de HDAC2 l'est par la SUMOylation des lysines 444 et 476 (David, Neptune et al. 2002). HDAC1 peut aussi être acétylée par p300, ce qui a un effet inhibiteur sur HDAC2 puisque ces 2 protéines ne peuvent alors plus s'hétérodimériser (Luo, Jian et al. 2009).

II.B.3 Désacétylation/acétylation des histones : un phénomène dynamique impliqué dans la régulation transcriptionnelle

II.B.3.a L'acétylation : une marque d'activation transcriptionnelle

II.B.3.a.i Corrélation entre acétylation et synthèse d'ARN

Une corrélation entre l'acétylation des histones et la synthèse d'ARN a été établie suite à des expériences d'incorporation de nucléotides ATP et UTP radioactifs pendant la synthèse d'ARN. La synthèse d'ARN est bloquée en présence d'histones totales isolées de thymus de veau et, plus la fraction d'histones est acétylée plus ce blocage est réduit (Allfrey, Faulkner et al. 1964). Des cinétiques d'incorporation d'acétate dans des lymphocytes (Pogo, Allfrey et al. 1966) ou des granulocytes de cheval (Pogo, Allfrey et al. 1967) traités avec un agent mitotique établissent que le taux d'acétylation des histones conditionne le taux de synthèse d'ARN et pas l'inverse (Gallwitz and Sekeris 1969). Ces résultats ont été également retrouvés *in vivo* suite à l'injection d'acétate marqué dans du foie en régénération de rat (1968 et 1969). Cette corrélation positive entre l'acétylation des histones et la synthèse d'ARN est observée chez les plantes, les animaux supérieurs mais fait l'objet d'un débat concernant les insectes comme la *Drosophila melanogaster* (Ellgaard 1967; Allfrey, Pogo et al. 1968).

Ces expériences indiquent donc un lien indirect entre l'acétylation des histones et l'activation transcriptionnelle.

II.B.3.a.ii Corrélation entre acétylation et activation transcriptionnelle

Après ces découvertes, il a été émis l'hypothèse qu'un lien direct existait entre l'acétylation des histones et l'activation transcriptionnelle. Certaines études ont permis d'appuyer cette idée (Sealy and Chalkley 1978; Jeppesen and Turner 1993; Grunstein 1997). Ainsi, une digestion de chromatine avec la DNase I, clivant préférentiellement la chromatine décondensée, a permis d'établir que les fractions les plus sensibles étaient enrichies en histones acétylées (Sealy and Chalkley 1978). Des expériences d'immunofluorescence menées sur des chromosomes humains ou murins en métaphase ont révélé que les anticorps dirigés contre l'acétylation de l'histone H4 se fixaient surtout dans des régions transcriptionnellement

actives alors que l'hypoacétylation concernait plutôt les régions inactives comme le chromosome X et l'hétérochromatine péricentromérique (Jeppesen and Turner 1993).

Puis, la découverte d'HAT a permis de corroborer cette hypothèse (Wolffe and Hayes 1999; Sterner and Berger 2000; Roth, Denu et al. 2001; Schones, Cui et al. 2008).

Les découvertes technologiques comme l'immunoprécipitation de chromatine couplée à un séquençage haut débit (ChIP-seq) ont ensuite permis d'observer à l'échelle du génome les sites ciblés par les HAT et les marques d'acétylation sur une multitude de sites. Des cartes de distribution de 5 HAT (Visel, Blow et al. 2009; Wang, Zang et al. 2009) ont été dressées et comparées à celles concernant l'acétylation sur 18 lysines et aux profils d'expression génique (Wang, Zang et al. 2008). Il a pu être confirmé par exemple que l'HAT p300 se retrouvait plutôt sur les promoteurs de gènes donc qu'elle participait à l'initiation de la transcription. Les marques d'acétylation sur les lysines 5 et 9 de l'histone H2A, 5, 12, 20 et 120 de l'histone H2B, 4, 9, 14 18, 23, 27, 36 de l'histone H3 et 5, 8 et 12 de l'histone H4 paraissent toutes associées à l'activation transcriptionnelle, ce qui avait déjà été établi pour la plupart de ces marques, et à la présence de HAT.

Aussi, une corrélation positive entre les HAT, ces modifications post-traductionnelles et la transcription génique a pu être établie sur une multitude de sites génomiques.

II.B.3.b La désacétylation, une marque de répression transcriptionnelle

II.B.3.b.i Recrutement des HDAC par les protéines liant l'ADN méthylé

Comme nous l'avons vu au cours du chapitre 2, les HDAC1 et HDAC2, présentes au sein de complexes remodelant la chromatine, sont recrutées par les protéines à MBD MeCP2 et MBD2 alors que la protéine HDAC3 est la seule à l'être par la protéine MBD1. Ces différences pourraient s'expliquer par les différences peptidiques observées au sein des protéines HDAC (voir paragraphes III.A.2.b.i et III.A.2.b.iii).

II.B.3.b.ii Quand les DNMT s'en mêlent

Les protéines DNMT catalysent la méthylation de l'ADN. Cependant, elles possèdent également toutes un domaine de répression transcriptionnelle qui pourrait leur permettre de jouer un autre rôle dans ce processus. Cette hypothèse est renforcée par le fait que ce domaine de la protéine Dnmt1 présente une homologie de séquence avec la protéine HRX impliquée dans la répression transcriptionnelle. Il a ainsi été révélé *in vivo* par la réalisation de chromatographies d'affinité et de co-immunoprécipitations que Dnmt1 recrute via ce domaine l'HDAC1, participant ainsi à la répression transcriptionnelle (Fuks, Burgers et al. 2000). La protéine Dnmt3b présente aussi ce rôle *in vivo* via le recrutement de Hdac2 pendant la différenciation neuronale dépendante de NGF (Bai, Ghoshal et al. 2005).

Ainsi, les protéines DNMT paraissent impliquées dans l'établissement du signal de méthylation d'ADN ainsi que dans sa traduction. Il est à noter que les protéines DNMT lient l'ADN non méthylé ce qui laisse supposer qu'elles induisent au moins une répression transcriptionnelle HDAC dépendante dans un contexte de non méthylation.

II.B.3.b.iii Corrélation entre désacétylation et répression transcriptionnelle

Nous avons vu au cours du chapitre 2 que les HDAC sont recrutées par les protéines à MBD afin d'initier la répression transcriptionnelle. De plus, au cours des mêmes expériences de CHIP-seq menées pour les HAT et l'acétylation sur les lysines (paragraphe III.A.3.a.ii), des cartes de distribution de 4 HDAC (HDAC1,2,3 et 6) ont été effectuées (Wang, Zang et al. 2009). De façon surprenante, elles semblent se trouver en majorité sur des promoteurs de gènes actifs pour HDAC1 et 3 et également sur leur région codante pour HDAC2. Leur présence corrèle avec les taux d'ARNm, d'ARN polymérase II et avec l'acétylation. Elles paraissent requises pour exciser le groupement acétyle après l'activation génique pour que la chromatine reprenne une forme compactée.

II.B.3.c HDAC et HAT : rôles spécifiques et complémentaires

II.B.3.c.i HDAC et HAT pour une même lysine ?

Etant donné que l'acétylation/désacétylation des histones correspond à un phénomène dynamique (Wade, Pruss et al. 1997; Eberharter and Becker 2002), il existe probablement pour chaque lysine une ou plusieurs HAT et HDAC qui remplissent des rôles spécifiques opposés. Ainsi, GCN5 peut acétyler le site H3K9 alors que HDAC1 le désacétyle pour qu'il puisse ensuite être méthylé. Récemment, il a été révélé que le site H3K56 est acétylé par p300 et désacétylé par SIRT2 et 3 *in vivo* (Yuan, Pu et al. 2009). Cette acétylation est retrouvée au niveau de gènes impliqués dans la tumorigenèse notamment dans la réponse aux dommages de l'ADN (Vempati, Jayani et al.). Cette modification apparaît souvent présente au niveau de gènes transcriptionnellement hyperactivés, aussi il reste à déterminer s'il existe une corrélation entre l'acétylation de H3K56 et l'activation transcriptionnelle.

Si les découvertes de HAT et HDAC se succèdent, il n'est pas encore déterminé pour tous les sites la ou les HAT et HDAC correspondantes.

II.B.3.c.ii Coopérations, compétitions ou régulations entre HDAC et HAT

Le rôle des HAT et HDAC ne s'arrête pas à une simple acétylation/désacétylation des histones pour la régulation transcriptionnelle.

En effet, il a été vu précédemment que p300 peut acétyler HDAC1 ce qui a pour effet d'empêcher son hétérodimérisation avec HDAC2 qui peut alors perdre son activité acétyltransférase. L'HAT a donc ici un rôle de régulateur d'activité HDAC.

De plus, il apparaît que HDAC3 et l'HAT p300 peuvent être recrutés par YY1 et interagir pour induire la répression transcriptionnelle du gène *c-myc* via la désacétylation au niveau du promoteur (Sankar, Baluchamy et al. 2008). Ceci suppose donc que HDAC et HAT coopèrent et que HDAC3 a un rôle direct alors que p300 intervient indirectement. Ces mêmes HAT et HDAC peuvent aussi entrer en compétition. En effet, KLF4 (Kruppel like factor 4) peut s'associer avec HDAC3 pour participer à la répression transcriptionnelle ou interagir

avec p300 pour être activé par acétylation et induire à son tour l'acétylation de l'histone H4, initiant l'activation transcriptionnelle (Evans, Zhang et al. 2007).

Ces HAT et HDAC peuvent donc intervenir en tant que co-facteurs de façon coopérative ou mutuellement exclusive et jouer un rôle direct ou indirect dans la régulation transcriptionnelle en acétylant/désacétylant les histones.

II.C La méthylation des histones

II.C.1 Une modification sur des arginines et lysines

Si la méthylation sur des protéines peut se retrouver au niveau des résidus lysines (K) (Murray en 1964), arginines (R) et histidines (H), il apparaît que seuls les sites K et R des histones peuvent être associés à la régulation transcriptionnelle.

Les arginines sont uniquement diméthylées alors que les lysines peuvent être mono-, di- ou triméthylées. Les enzymes capables d'une activité méthyltransférase sur des histones sont les protéines de la famille Trithorax et les HMT.

II.C.2 Une modification catalysée par des HMT

Les méthyltransférases d'histones (HMT) catalysent le transfert de groupement(s) méthyle(s) sur le groupe amine à partir du substrat S-adénosyl-méthionine.

Ces HMT sont réparties en fonction du site de catalyse. La méthylation des arginines est effectuée par la famille de PRMT (Protein arginine methyltransferase) de type I si la diméthylation est asymétrique et II si elle est symétrique. La méthylation des lysines concerne surtout la famille à domaine SET (Su(var)3-9, Enhancer of Zeste, Trithorax) puis DOT1L (Dot1-like) ou KMT4 qui pourrait être le membre fondateur d'une famille de HMT sans domaine SET (Feng, Wang et al. 2002). A titre d'exemples, PRMT5 apporte 2 groupements méthyles sur les sites R8 de l'histone H3 (Pal, Vishwanath et al. 2004). Parmi les membres de la famille SET, Smyd2 catalyse la di-méthylation de H3K36 (Brown, Sims et al. 2006), l'HMT WHISTLE la di- et tri-méthylation de H3K27 et la diméthylation de H3K4 (Kim, Kee et al. 2006). DOT1L peut mono-, di- ou tri-méthyle H3K79 (Steger, Lefterova et al. 2008).

La mono-, di- ou tri-méthylation d'une lysine peut faire intervenir différentes enzymes et/ou dépendre du domaine chromatinien. Ainsi, la mono-méthylation de H4K20 est catalysée par PR-SET7 alors que sa tri-méthylation concerne SUV420H1. G9a est responsable de la monométhylation de H3K9 dans l'euchromatine, de la diméthylation dans l'hétérochromatine, alors que sa triméthylation est effectuée par Suv39h1 and Suv39h2 (Rice, Briggs et al. 2003). Les enzymes SETDB1/ESET et GLP/EuHMTase1 interviennent dans sa diméthylation. Pas moins de 5 HMT peuvent donc contribuer à la méthylation de ce site.

La régulation de ces HMT peut également se faire par des modifications post-traductionnelles comme la phosphorylation sur CARM1, une PRMT de type I (Feng, He et al. 2009). Des protéines sont aussi capables d'interagir avec elles pour contrôler leur activité de transfert de méthylation comme les protéines RB1 avec les HMT Suv4-20h1 et 2 catalysant la triméthylation de H4K20 (Gonzalo, Garcia-Cao et al. 2005).

II.C.3 La méthylation des histones, une marque associée à l'activation ou à la répression transcriptionnelle

La découverte de la méthylation sur les résidus lysines et arginines des histones ont apporté de nombreuses interrogations quant à leur(s) rôle(s) surtout dans la régulation transcriptionnelle. Il apparaît que la méthylation des histones est une marque d'activation ou de répression transcriptionnelle selon le résidu lysine touché, le nombre de groupements méthyles ajoutés ainsi que les domaines chromatiniens impliqués.

La méthylation des histones peut être associée à une répression transcriptionnelle mais ce n'est pas corrélé à un nombre de groupements méthyles particulier. En effet, les di- (Rice, Briggs et al. 2003) ou triméthylation de la lysine 9 de l'histone H3 (H3K9) (Rea, Eisenhaber et al. 2000; Nakayama, Rice et al. 2001; Peters, Mermoud et al. 2002) sont associées à la répression transcriptionnelle alors que sa monométhylation est plutôt retrouvée sur des promoteurs activés (Barski, Cuddapah et al. 2007). L'association de Suv39H, responsable de cette triméthylation, avec HP1, une protéine à chromodomaine, paraît nécessaire pour la formation et la propagation de l'hétérochromatine et donc la répression génique (Nielsen, Schneider et al. 2001; Shi, Sawada et al. 2003). La méthylation en H3K9 ne peut avoir lieu

que si le groupement a été précédemment désacétylé (Nicolas, Roumillac et al. 2003) puisque méthylation et acétylation sont des marques mutuellement exclusives sur un même site.

D'autres méthylations comme les triméthylations retrouvées sur les lysines 27 de l'histone H3 (Rougeulle, Chaumeil et al. 2004) et 20 de l'histone H4 (Schotta, Lachner et al. 2004) corrélient aussi avec la répression transcriptionnelle et se situent dans l'hétérochromatine. Il existe des corrélations entre tous ces sites méthylés. En effet, la di-méthylation en H3K9 et la tri-méthylation en H3K27 sur le chromosome X inactif semblent avoir la même cinétique d'apparition mais être complémentaires (Rougeulle, Chaumeil et al. 2004). La triméthylation de H3K9 précède et paraît nécessaire à celle de H4K20 (Schotta, Lachner et al. 2004).

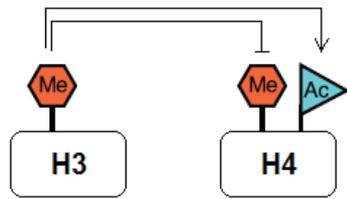
La méthylation des histones répond à un mécanisme complexe puisqu'elle peut aussi être associée à l'activation transcriptionnelle en modifiant le nombre de groupements méthyles sur les sites associés à une répression transcriptionnelle ou en touchant d'autres sites. Ainsi, la tri-méthylation de la lysine 4 de l'histone H3 est associée à l'activation transcriptionnelle (Liang, Lin et al. 2004; Schneider, Bannister et al. 2004; Heintzman, Stuart et al. 2007) alors que des expériences menées chez *S.cerevisiae* laissent supposer que sa di-méthylation est associée à l'activation et à la répression transcriptionnelle (Santos-Rosa, Schneider et al. 2002). Il paraît que SMYD3 qui catalyse la méthylation sur le site H3K4 appartient à un complexe impliqué dans la machinerie transcriptionnelle (Hamamoto, Furukawa et al. 2004).

D'autres méthylations comme la di- ou triméthylation de la lysine 79 (Steger, Lefterova et al. 2008) et la triméthylation de la lysine 36 de l'histone H3 (Bernstein, Kamal et al. 2005), la monométhylation de H3K20 (Barski, Cuddapah et al. 2007) sont également associées à l'activation transcriptionnelle et à l'élongation ainsi que la méthylation de l'arginine 3 de l'histone H4 (Wang, Huang et al. 2001).

II.C.4 Les déméthylases

Les découvertes de résidus méthylés sur les histones et de HMT laissaient supposer que la méthylation correspondait à un phénomène stable alors que le parallèle fait avec la caractérisation du phénomène dynamique d'acétylation/désacétylation tendait plutôt à inscrire la méthylation dans une dynamique (Rice and Allis 2001). La découverte de

A. Interactivité des marques entre histones



acétylation

méthylation associée à répression

méthylation associée à activation

B. Interactivité des marques au sein d'une histone

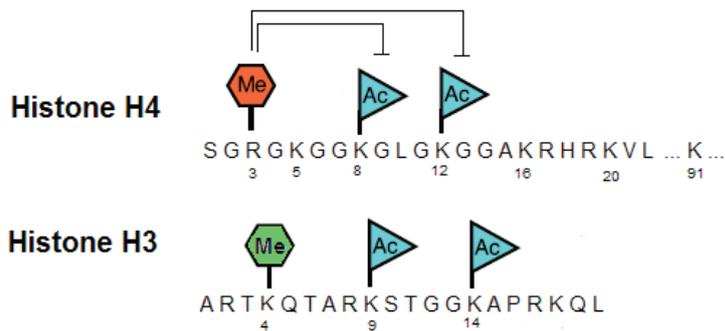


Figure 22 : Interactivités des marques d'acétylation et de méthylation selon leurs positions sur les histones H3 et H4

méthyltransférases d'histones, initiée par les travaux de Paik et de ses collaborateurs, a permis de trancher et d'établir l'existence d'un phénomène dynamique de méthylation/déméthylation.

La première déméthylase identifiée fut LSD1 (Lysine Specific Demethylase 1) par des expériences similaires à celles ayant permis de caractériser GCN5 (Shi, Lan et al. 2004). Elle catalyse la déméthylation de 2 groupements méthyles sur la lysine K4 de l'histone H3.

Il existe des déméthylases qui déméthylent par un processus oxydatif : les JHDM (JmjC domain-containing histone demethylase). Ainsi, JHDM3A, JMDM2A et la famille des JMJD2 (jumonji domain containing 2) peuvent exciser la triméthylation présente sur les sites H3K9 et H3K36 pour aboutir à une mono-méthylation (Klose, Yamane et al. 2006; Whetstine, Nottke et al. 2006; Yamane, Toumazou et al. 2006). JHDM1 peut prendre le relais de l'une ou l'autre puisqu'elle excise la mono- ou di-méthylation de H3K36 (Tsukada, Fang et al. 2006).

II.D Combinaison de méthylation et d'acétylation sur les histones

Il existe un lien entre les marques d'acétylation et de méthylation présentes sur les mêmes histones ou entre différentes histones. Il a été révélé que la méthylation de l'histone H4 s'effectue plutôt s'il n'y a pas d'acétylation déjà présente sur cette histone alors que c'est l'opposé pour l'histone H3. De plus, la méthylation en H3 est corrélée à la présence d'acétylation et à la perte de méthylation en H4 (Liu, Teng et al.; Annunziato, Eason et al. 1995) (**Figure 22A**). Ses résultats ont été corroborés et précisés. En effet, la méthylation du site H4R3 par PRMT1 favoriserait l'acétylation des sites H4K8 et H4K12 par l'HAT p300 alors qu'à l'inverse, l'acétylation de ces sites inhibe la méthylation du site H4R3 (Wang, Huang et al. 2001) (**Figure 22B**). L'HMT Smyd2 paraît avoir besoin du recrutement d'un complexe désacétylase sur l'histone H3 en K36 avant de dimétyler ce site (Brown, Sims et al. 2006).

Si ce sont des marques mutuellement exclusives sur un même site, elles peuvent néanmoins co-exister sur les mêmes histones. En effet, la méthylation en H3K4 peut se retrouver localiser sur des régions situées en 5' de promoteurs de gènes transcrits au même titre que l'acétylation en H3K9 et H3K14 (Liang, Lin et al. 2004) (**Figure 22B**).

Toutes ces données sont en accord avec l'existence d'un code des histones qui suppose une combinaison des modifications post-traductionnelles des histones impliquées dans la régulation transcriptionnelle. Aussi, récemment, ces marques sont étudiées en combinaison

(Bonenfant, Towbin et al. 2007; Winter, Fischle et al. 2008), pour des loci particuliers (Tian, Hu et al. 2009) ou pour tout le génome par des approches globales (Wang, Zang et al. 2008).

La combinaison de ces marques passe par la coopération d'enzymes responsables de l'ajout ou de l'excision des groupements acétyles ou méthyles. Des études ont révélé l'existence d'un complexe SMCX contenant des HMT (agissant sur deux sites différents) et des HDAC (Tahiliani, Mei et al. 2007) ou d'un complexe formé par l'HMT Suv39H1-HP1 recruté par MBD1 et recrutant à son tour HDAC1 et 2 (Fujita, Watanabe et al. 2003) d'où une coopération nécessaire pour enlever la marque d'acétylation avant d'ajouter des groupements méthyles. De plus, il a pu être mis en évidence une association entre la méthyltransférase d'histones CARM1 ou PRMT4, et l'acétyltransférase d'histones p300 qui permettrait une synergie d'action avec le facteur nucléaire KappaB (NF-KappaB) afin réguler l'activation transcriptionnelle *in vivo* de gènes cibles de NF-KappaB (Covic, Hassa et al. 2005).

Des technologies pour des études à grande échelle du signal de méthylation de l'ADN, de ses lecteurs et des changements chromatinien

I- Cartographies des sites méthylés de l'ADN

Différentes méthodes existent pour étudier la méthylation de l'ADN sur des régions particulières. Nous allons voir comment elles ont été adaptées afin de rendre leurs utilisations moins lourdes et coûteuses pour une étude globale du génome.

I.A A partir de digestions enzymatiques

La première technique permettant d'étudier la méthylation de l'ADN a été la digestion enzymatique. Elle consiste à employer une enzyme de restriction, d'origine bactérienne, comme *NotI* ou *HpaII* (Bird, Taggart et al. 1985) sensible à la méthylation au niveau de son site de reconnaissance. Un clivage sera observé si les CpG sont non méthylés.

Pour une séquence particulière, une amplification de cette région est réalisée par PCR puis les amplicons sont digérés et séparés sur gel d'électrophorèse.

Pour étudier globalement le génome, une première méthode combine 2 digestions enzymatiques et une migration bidirectionnelle sur gel. Une enzyme sensible à la méthylation coupe l'ADN total, générant des fragments qui seront marqués à leurs extrémités puis séparés par électrophorèse avant l'action d'une 2^e enzyme coupant fréquemment le génome et suivie d'une 2^e migration dans un sens différent. Mais cette technique appelée RLGS (restriction landmark genomic scanning) ne sert qu'à comparer des échantillons. Aussi, dans une autre variante, l'ADN est digéré en inversant l'ordre de ces enzymes et le marquage est remplacé par une ligation de petits ADN servant à l'amplification avant un dépôt sur puces à ADN. Or,

toutes ces méthodes permettent une identification indirecte des sites méthylés par absence de coupure. Aussi, sur une même puce sont hybridés des échantillons digérés, soit avec *HpaII* soit avec *MspI* reconnaissant le même site mais insensible à la méthylation, et marqués par des fluorochromes différents. Ainsi, cette méthode, appelée HELP (*HpaII* tiny fragment enrichment by ligation-mediated PCR) permet de connaître par un signal direct la présence ou l'absence de méthylation sur tous les CpG digérés (Khulan, Thompson et al. 2006).

I.B A partir de modifications chimiques par bisulfite

L'utilisation de bisulfite de sodium (NaHSO_3) permet de modifier chimiquement les cytosines non méthylées en des uraciles alors que les cytosines méthylées sont protégées de ce changement. L'analyse des échantillons modifiés peut se faire par MSP (methylation specific PCR) en employant un double jeu d'amorces avec une paire spécifique de la méthylation et une paire de la non-méthylation et contenant toutes un ou plusieurs sites CpG à discriminer.

Pour couvrir plus de CpG, d'autres techniques sont employées. La MethyLight quantitative PCR (Wild, Funes et al.) consiste, à partir d'amorces situées en dehors de sites CpG, à amplifier l'ADN, souvent à cloner les amplicons avant de les séquencer afin d'analyser les CpG présents à l'intérieur (Chatagnon, Ballestar et al.; Hodge, Peng et al. 2005). Il est à noter qu'au cours de l'amplification, les uraciles sont remplacés par des thymines. Les techniques de séquençage à haut débit suppriment l'étape de clonage et permettent l'analyse simultanée de nombreuses séquences. Parmi ces méthodes, le pyroséquençage consiste à ajouter les uns après les autres les nucléotides marqués. Si le nucléotide ajouté est incorporé dans la séquence en cours d'élongation, il y a libération d'un pyrophosphate qui va alors être transformé pour aboutir au final à un signal lumineux qui sera capté par les caméras du séquenceur. La proportion de C par rapport au T pour chaque position reflète le pourcentage de brins méthylés ou déméthylés sur ce site dans l'échantillon initial. Les limites principales de cette technique sont de ne pas distinguer le profil des deux allèles et de ne pouvoir analyser que des courtes séquences (<100pb).

I.B.1 A partir de précipitations d'ADN ou de chromatine

L'ADN simple brin peut être précipité par des anticorps dirigés contre la 5-méthylcytosine. Une analyse par PCR permet ensuite de comparer la fraction précipitée à la fraction initiale. Un enrichissement traduit la présence de CpG méthylés. L'hybridation sur différentes puces à ADN des fractions précipitées et initiales permet d'analyser globalement le génome (Weber, Hellmann et al. 2007). Cette technique correspond au ChIP-on-chip.

Cependant, ce type d'anticorps reconnaît préférentiellement les sites riches en CpG méthylés. Aussi, pour éviter des biais dus à la densité en CpG, l'ADN double brin est précipité par des protéines chimériques recombinantes composées essentiellement de domaines de liaison à l'ADN méthylé spécifiques des protéines à MBD (MBD2 et MeCP2) (Serre, Lee et al.; Cross, Charlton et al. 1994) ou retenu sur des colonnes recouvertes de domaines de liaison à l'ADN non méthylé CXXC présent dans MBD1 ou Cfp1 (Thomson, Skene et al.; Illingworth, Kerr et al. 2008). L'analyse de ces précipitations permet respectivement de distinguer les régions présentant des CpG méthylés et non méthylés. Les fractions issues de ces méthodes sont analysables de la même façon que celles issues des immunoprécipitations de chromatine (ChIP) ou alors par séquençage haut débit, comme pour les fractions issues de l'IP avec des protéines recombinantes contenant le MBD de MBD2 (Serre, Lee et al.).

I.B.2 A partir d'une combinaison de techniques

Il existe également des combinaisons de techniques afin de détecter rapidement et de façon quantitative les sites CpG méthylés. Ainsi, une digestion enzymatique avec des enzymes sensibles à la méthylation est associée à une précipitation d'ADN méthylé avec protéines chimériques contenant le domaine de liaison à l'ADN des protéines à MBD (Yegnasubramanian, Lin et al. 2006). Dans cette étude, l'analyse s'est faite par PCR quantitative mais il paraît envisageable de réaliser un séquençage à haut débit ou une hybridation sur puces pour une analyse globale du génome.

Cette méthode, aussi appelée COMPARE-MS, ne nécessite pas de modification au bisulfite de sodium, ce qui a l'avantage de préserver la séquence nucléotidique initiale. Elle serait également plus sensible que la technique de MethyLight, permettant ainsi de détecter plus de sites méthylés, et plus spécifique, diminuant le nombre de faux positifs.

II- Cartographies des sites de fixation des protéines

Deux méthodes sont utilisées pour identifier directement les sites de fixation des protéines *in vivo* : les DamID et ChIP.

II.A Modification d'ADN par DAM fusionnée

La méthyltransférase d'adénines issue d'*E.coli* (DAM) méthyle les adénines présentes dans les motifs GATC. Cette propriété va être utilisée dans la méthode DamIP. Elle consiste à fusionner une protéine d'intérêt avec Dam et à l'exprimer dans la cellule cible. Dam va alors méthyle les sites GATC présents à proximité des sites fixés par la protéine d'intérêt. Puis, l'ADN génomique est isolé et digéré sélectivement avec l'enzyme de restriction *DpnI* qui reconnaît le motif palindromique GATC méthylé en A. Des adaptateurs aux extrémités des fragments coupés sont ajoutés avant de digérer l'ADN avec *DpnII* qui va couper au niveau du motif GATC non méthylé en A. Puis une amplification a lieu avant hybridation sur puces à ADN (Vogel, Peric-Hupkes et al. 2007). L'inconvénient majeur de cette technique est d'engendrer des faux positifs à cause de la surexpression de la protéine de fusion.

II.B Immunoprécipitation de chromatine avec des anticorps

L'immunoprécipitation de chromatine consiste à lyser les cellules et à fragmenter la chromatine par digestion avec une nucléase micrococcalle pour la chromatine native (plutôt employée pour l'étude des histones liées de façon stable à l'ADN) ou par sonication (employée pour les facteurs de transcription et les protéines liées à la chromatine) si les protéines ont été liés chimiquement de façon covalente à leurs sites de fixation sur l'ADN, par des agents de « pontage » comme le formaldéhyde. Puis une précipitation est réalisée par des anticorps dirigés contre la protéine d'intérêt. La spécificité de cet anticorps détermine celle du ChIP. L'analyse se fait par hybridation sur puces à ADN avec une éventuelle pré-amplification (Wang, Zang et al. 2009) ou par séquençage à haut débit (Skene, Illingworth et al.) appelé ChIP-seq. On parle de ChIP-PET (ChIP Paired End diTags) dans le cas particulier où des ligands sont ajoutés aux extrémités des fragments ADN avant le séquençage (Wang,

Objectifs de travail

Xu et al.). Il existe même une technique, appelée ChIA-PET, permettant d'identifier les sites de liaison de la protéine d'intérêt tout en conservant la structure en 3D de la chromatine initiale. Cette approche est donc une combinaison de ChIP-PET et de « Chromosome Conformation Capture » (3C) (Fullwood, Liu et al. 2009) qui présente l'avantage de réduire la complexification de la 3C et d'augmenter la spécificité du ChIP-PET.

III- Une cartographie des modifications des histones

De la même façon que pour identifier les sites de fixation des protéines à l'ADN, des immunoprécipitations de chromatine (ChIP) avec des anticorps reconnaissant spécifiquement une histone et sa modification sont effectuées (Wang, Zang et al. 2008). L'utilisation de la nucléase micrococcale permet de digérer de façon ménagée la chromatine native entre les nucléosomes. Ainsi, contrairement aux ChIP pour lesquels un agent chimique est employé pour lier les protéines à l'ADN, les nucléosomes restent intacts et la conformation des épitopes reconnus par les anticorps est conservée, ce qui augmente la spécificité de l'anticorps.

L'analyse se fait aussi par hybridation sur puces ou séquençage haut débit.

IV- Un choix de puces à ADN déterminant

Presque toutes les techniques décrites dans ce chapitre peuvent se terminer par une étape d'hybridation sur puces à ADN, ce qui permet de cartographier les sites de méthylation de l'ADN, les sites de fixation des protéines liant l'ADN ou les histones et leurs modifications. Le choix de la puce d'ADN est un facteur déterminant selon les régions qui suscitent notre intérêt, ceci afin de limiter le nombre considérable de données générées suite à l'analyse et/ou comparaison des puces. Ainsi, si tout le génome peut être étudié, d'autres puces restreignent les zones observées comme les puces dédiées aux promoteurs, aux régions riches en CpG ou aux exons.

Nous avons vu au cours de la revue bibliographique que chez les mammifères, la méthylation de l'ADN concerne les cytosines présentes au niveau de séquences symétriques 5' C-G 3' ou doublets CpG, dispersées au sein du génome et exceptionnellement, des régions riches en CpG ou îlots de CpG. Or, de nombreux travaux ont montré une altération des profils de méthylation dans les cancers. En effet, une hypométhylation globale du génome entraînant l'expression ectopique de certains gènes et une instabilité génomique est associée à une hyperméthylation de certains îlots de CpG, induisant la répression de certains gènes suppresseurs de tumeurs. De plus, il est maintenant établi que les doublets CpG méthylés peuvent être notamment reconnus par une famille de protéines possédant le même motif « Methyl-CpG-binding domain » ou MBD.

Une relation directe entre la fixation des protéines MeCP2, MBD1 et MBD2, appartenant à cette famille, et la répression de gènes suppresseurs de tumeurs a été démontrée. Ces résultats émanent très majoritairement d'études réalisées à partir de lignées cellulaires et tissus cancéreux. Or, peu d'études prennent pour modèles des lignées cellulaires et/ou tissus normaux et il manque souvent la comparaison entre les modèles cancéreux et leurs contreparties normales. De plus, la distribution de ces protéines n'est a priori définie pour aucun modèle de progression tumorale.

Enfin, nous avons vu que la protéine MBD2 semble au moins pour certains cancers impliquée dans le processus tumoral et que son absence touche le plus de gènes par rapport aux autres répresseurs MeCP2 et MBD1.

Aussi, mes objectifs sont d'identifier:

- **La mise en place des modifications des profils de méthylation de l'ADN dans un modèle homogène de progression tumorale**
- **La distribution de la protéine MBD2 dans un modèle cellulaire normal ; cancéreux ; au cours de la progression tumorale**
- Rôle joué par la méthylation de l'ADN dans cette distribution
- **Rôle(s) joué(s) par la protéine MBD2 dans ces modèles par rapport à la régulation de l'expression génique**
- Impact de l'absence de la protéine MBD2 sur l'expression génique
- Rôle(s) joué(s) par la protéine MBD2 modifié(s) en fonction de la distribution de la protéine MBD2 ?

En réponse à ces interrogations, nous avons déjà réalisé une approche « gène candidat » au cours de laquelle nous avons montré que la **protéine MBD2 peut se fixer de façon spécifique et exclusive sur le promoteur du gène *hTERT* s'il est méthylé** et induire une répression transcriptionnelle. De plus, nous avons mis en évidence, par interférence à l'ARN transitoire ou constitutive, que **l'absence de MBD2 peut participer à la réexpression génique (article 1, en annexe 1)**.

Mais il paraît incontournable d'élargir ce type d'analyse à l'ensemble des promoteurs du génome pour **identifier les cibles spécifiques de MBD2 dans différents modèles cellulaires normaux, cancéreux et de progression tumorale** en développant des **approches globales**. Ainsi, cela permettrait notamment de **comprendre quand et comment la protéine MBD2 ségrège entre les sites méthylés, au cours de la progression tumorale** afin de déterminer si **l'altération des profils de méthylation de l'ADN entraîne directement une modification des profils de liaison de la protéine MBD2 et quelles répercussions ont toutes ces modifications sur les changements d'expression génique** observés dans les cancers.

De plus, cela permettra d'évaluer si le ciblage de MBD2 peut représenter une nouvelle thérapie épigénétique plus sélective, en alternative aux traitements épigénétiques actuels comme la 5-aza-2' desoxycytidine qui touche les enzymes catalysant la méthylation de l'ADN mais pourrait conduire à une instabilité chromosomique.

Résultats

Article 2

*Relationship Between the CpG Content of Gene Promoter and
Transcriptional Repression by the Methyl-CpG Binding Domain
Protein 2 in Normal Human Fibroblastic Cells*

Laury Perriaud, Joel Lachuer and Robert Dante

Soumis à *Epigenetics*

Contexte d'étude

La plupart des sites de liaison de la protéine MBD2 ont été identifiés dans les lignées cellulaires et tissus cancéreux. Il apparaît que la majorité de ces cibles réprimées entre dans la catégorie des gènes suppresseurs de tumeurs présentant des promoteurs inclus dans des îlots de CpG hyperméthylés dans les cancers. Mais quelles sont les cibles de la protéine MBD2 dans les cellules normales qui présentent un profil de méthylation différent ? Cette protéine joue-t-elle un rôle important dans la répression transcriptionnelle étant donné que les souris *Mbd2*^{-/-} présentent un phénotype très peu altéré ?

Stratégie de l'étude

Il a été récemment établi dans des fibroblastes pulmonaires primaires humains (WI-38) un classement de promoteurs en 3 catégories en fonction de leur densité en CpG et taux en C+G. Il apparaît que la méthylation de l'ADN semble très peu concerner les régions hautement riches en CpG. De plus, il semble que les promoteurs présentant une concentration plus faible mais quand même notable en CpG (promoteurs appelés ICP) sont la cible principale des changements de méthylation corrélés à l'expression génique au cours de la différenciation (Weber, Hellmann et al. 2007).

Aussi, nous avons choisi d'étudier dans un **modèle cellulaire analogue**, les **cellules normales fibroblastiques embryonnaires de poumon (MRC5)**, les **conséquences de l'inhibition de la protéine MBD2 suite à une interférence à l'ARN sur l'expression génique** et de déterminer **quelle(s) catégorie(s) de promoteur(s) étai(en)t majoritairement concerné(s)**.

Résultats

Après avoir **vérifié l'efficacité et la spécificité de l'interférence à l'ARN dirigé contre l'ARNm MBD2**, les conséquences de cette inhibition ont été évaluées en comparant par une approche globale les profils d'expression génique entre les cellules normales et invalidées, avec deux échantillons indépendants par condition. Pour chaque échantillon, 1µg d'ARN total extrait par kit Qiagen a été contrôlé pour son intégrité par la plateforme

ProfilExpert (Lyon) avant d'être rétrotranscrit en ADNc, marqué puis hybridé sur une puce d'expression CodeLink™ Human Whole genome, selon les protocoles préconisés par Affymetrix. Ces puces ont été choisies car elles permettent d'analyser l'expression d'environ 55.000 transcrits dont 45.000 bien caractérisés dans la base NCBI/Unigene et recouvrant les principaux processus biologiques. De plus, l'analyse globale est facilitée par le fait que chaque transcrit n'est représenté qu'environ 1 fois sur la puce (avec ~1sonde/transcrit). Au final, après n'avoir retenu que les résultats similaires entre les puces de même condition, **il apparaît que peu de gènes (571) sont touchés par l'inactivation de l'ARNm MBD2 car ils représentent <2% des gènes présents sur les puces.**

Sur ces 571 gènes, 178 sont bien annotés et sont gardés pour les analyses futures. Parmi ces derniers, **certains deviennent exprimés (42) alors que les autres (136) voient leur expression modulée** par rapport aux témoins. Ceci est en accord avec les dernières données de la littérature montrant que MBD2 pourrait avoir un rôle répresseur strict et parfois de modulateur de la transcription.

Avant de déterminer si ces 42 gènes représentent des cibles directes ou plutôt ou indirectes de la protéine MBD2, les expressions de 6 d'entre eux (*LIMS2*, *EGFL7*, *PRPH2*, *FAM65A*, *DNASE1*, et *IL2RB*) et d'un gène silencieux dans les deux conditions ont été quantifiées par RT-PCR dans d'autres expériences indépendantes afin de valider les résultats obtenus sur les puces CodeLink. Puis, le taux de méthylation et la présence de MBD2 au niveau de la région promotrice de ces 6 gènes ont été déterminés respectivement par bisulfite et séquençage et par immunoprécipitation de chromatine avec des anticorps dirigés contre MBD2. Il apparaît que **ces promoteurs sont méthylés et qu'ils sont tous liés par MBD2**, faisant d'eux des **cibles directes de la protéine MBD2 dans les cellules normales**. Aussi, ces résultats **suggèrent que MBD2 initie la répression méthyl-dépendante de la plupart de ces 42 gènes**, ce qui représenterait **au maximum ¼ des gènes altérés suite à l'interférence à l'ARN**.

En parallèle, les **groupes de gènes ré-exprimés et stimulés ont été caractérisés selon deux critères: leur composition nucléotidique et leur fonction moléculaire.**

Ainsi, la densité en CpG et le taux de C+G ont été déterminés pour ces 2 groupes avec le logiciel CpGProd au niveau de leur région promotrice comprise entre -900 et +400pb par rapport au site d'initiation de la transcription. Puis, une répartition de ces promoteurs a été

établie selon les critères établis par Weber et ses collaborateurs afin de réaliser une étude comparative avec leurs résultats obtenus dans la lignée WI-38. Si les 136 gènes présentent une distribution similaire à celle observée pour les promoteurs méthylés de WI-38, les gènes ré-exprimés présentent une distribution significativement différente avec une forte sur-représentation des gènes appartenant à la catégorie ICP (61%) et une forte sous-représentation des HCP (2 gènes). J'ai montré donc par une interférence à l'ARN qu'une **forte diminution des concentrations de MBD2 dans des cellules normales fibroblastiques embryonnaires de poumon (MRC5) stimulait très préférentiellement l'expression des gènes ayant des promoteurs ICP** et qu'au contraire, elle ne **semblait pas suffisante pour « déverrouiller » les gènes présentant des HCP.**

De plus, chacune de ces 2 listes a été soumise sur le site internet GOTM afin de déterminer les fonctions moléculaires majeures occupées par ces 2 groupes. Il apparaît ainsi que les **42 gènes occupent un rôle essentiel au sein de la machinerie transcriptionnelle** alors que les **136 gènes semblent plus impliqués dans les voies métaboliques.**

Ainsi, j'ai pu montrer qu'une **inhibition de la protéine MBD2 suite à une interférence à l'ARN altérait l'expression de peu de gènes dans des cellules normales. Sur les 178 gènes, les 3/4, essentiellement impliqués dans les voie métaboliques, voient leur expression stimulée alors que 1/4 d'entre eux, surtout engagés dans la machinerie transcriptionnelle, sont ré-exprimés.** De plus, les résultats obtenus suggèrent que **dans des cellules normales, la protéine MBD2 intervient dans l'initiation de la répression transcriptionnelle majoritairement suite à sa fixation sur des promoteurs méthylés présentant une densité en CpG notable (promoteurs ICP).**

Relationship between Gene Promoter CpG Content and Transcriptional Repression by Methyl-CpG Binding Domain Protein 2 in the Normal Human Cell Line MRC5.

Laury Perriaud¹, Joël Lachuer² and Robert Dante^{1*}.

¹ INSERM, U590, Lyon, F-69008, France;

² ProfileXpert/Neurobiotec Service, INSERM U842, Bron, F-69676, France.

* Corresponding author: Dr.Robert Dante
Unité INSERM 590
Oncogenese et progression tumorale
Centre Léon Bérard
28 rue Laennec
69373 Lyon Cedex 08
FRANCE
Phone: 33(0)4-78-78-59-22
Fax: 33(0)4-78-78-27-20
Email dante@univ-lyon1.f

Key words: DNA methylation - Transcription - Methylation dependent repressor

ABSTRACT

Background.

Transcriptional repression of methylated genes is at least partly mediated by the MBD2 protein (methyl-CpG binding domain protein 2), through recruitment of chromatin remodeling complexes. Yet MBD2 depletion does not result in increased expression of all MBD2-target genes. To characterize the genes showing upregulated expression upon MBD2 depletion, we applied microarray analysis to the normal human embryonic fibroblast cell line MRC-5 to investigate the impact on gene expression of MBD2 inhibition by RNA interference.

Methodology/Principal Findings.

MBD2 depletion did not alter globally the levels of the 55,000 transcripts represented on the arrays. Among these, 536 transcripts were found to be upregulated. We chose to focus on 178 genes annotated with high confidence, placing them in two groups: genes expressed in control cells and whose expression was enhanced by MBD2 depletion (136 “depletion-stimulated” genes) and genes not expressed in control cells but whose expression was triggered in response to MBD2 depletion (42 “depletion-activated genes”). Gene ontology analysis revealed different molecular functions for these two groups. When the promoter regions of the two groups of upregulated genes were classified according to their CpG content, the 136 “depletion-stimulated” genes showed a promoter-class distribution resembling that of methylated genes throughout the genome, while the 42 “depletion-activated” genes exhibited a specific distribution of promoter classes (χ^2 test, $p = 10^{-4}$), with only two of them having CpG-rich promoters.

Conclusions/Significance.

Taken together, these data indicate that in normal fibroblasts, MBD2 depletion might not trigger expression of silent CpG-rich genes, although removal of this methylation-dependent

repressor can boost transcription of both CpG-rich and other genes expressed in its presence. This finding suggests that in normal fibroblasts, gene silencing mediated by CpG island methylation is associated with a tighter repressive chromatin structure than gene silencing at promoters having a low or intermediate CpG content.

ABSTRACT

Transcriptional repression of methylated genes is at least partly mediated by the MBD2 protein (methyl-CpG binding domain protein 2), through recruitment of chromatin remodeling complexes. To characterize the genes activated upon MBD2 depletion, we used microarray analysis to investigate, in the normal human embryonic fibroblast line MRC-5, the impact of MBD2 inhibition (by RNA interference) on gene expression. A gene was considered “expressed” if the corresponding signal on duplicate microarrays exceeded the average signal of a set of negative controls plus three standard deviations. MBD2 depletion did not lead to global gene expression changes. We identified 178 upregulated genes annotated with high confidence, placing them in two groups: “depletion-stimulated” genes (136), which were expressed in control cells and whose expression was further enhanced in MBD2-depleted cells, and “depletion-activated” genes (42), classified as silent in control cells but not in MBD2-depleted cells. Gene ontology analysis revealed different molecular functions for these two groups. When the promoter regions of both groups of upregulated genes were classified according to their CpG content, the 136 “depletion-stimulated” genes showed a promoter-class distribution resembling that of methylated genes throughout the genome, while the 42 “depletion-activated” genes exhibited a specific distribution of promoter classes (χ^2 test, $p = 10^{-4}$), with only two of them having CpG-rich promoters.

When the promoter regions of the two groups of upregulated genes were classified according

to their CpG content, the 136 “depletion-stimulated” genes showed a promoter-class distribution resembling that of methylated genes throughout the genome, while the 42 “depletion-activated” genes exhibited a specific distribution of promoter classes (χ^2 test, $p = 10^{-4}$), with only 2 out of 42 having CpG-rich promoters. This finding suggests that in normal fibroblasts, gene silencing mediated by CpG island methylation is associated with a tighter repressive chromatin structure than gene silencing at promoters having a low or intermediate CpG content.

INTRODUCTION

The correlation between DNA methylation and gene silencing has been recognized for many years. Although direct inhibition of transcription factor binding can occur, it is now widely accepted that methylation marks need to be translated by proteins that recognize methylated cytosines at 5'-C-G-3' (CpG) sequences [1]. These methyl-DNA binding proteins (MBP) belong to three distinct protein families [2]: the methyl-CpG binding domain (MBD) proteins (MeCP2, MBD1/2/3/4), the SRA proteins (SET and Ring finger associated, UHRF1/2), and the methyl-DNA binding zinc finger proteins (Kaiso, ZBTB4/38). The MBD proteins were the founding members of the MBP families. They provide a direct link between DNA methylation and repressive chromatin structure mediated by histone modifications [3].

In several cancer cell lines, MBD2 associates with many silent genes exhibiting aberrantly hypermethylated CpG islands (CGi) [4-12]. In a uterine cervix carcinoma cell line, MBD2 depletion was shown, on arrays representing a limited number of genes, to be more often involved in reactivating hypermethylated tumor suppressor genes than either MeCP2 or MBD1 [10].

Few studies have focused on the repressive role of Mbd2 in normal cells, and most of the available data concern mice, usually genetically modified [13-18]. In transiently transfected tail fibroblasts from *Mbd2*^{-/-} mice, methylation-dependent repression of reporter genes linked to in vitro methylated promoters was shown to be impaired [19]. In contrast, *Mecp2* knockout did not affect this repression [19]. This suggests that Mbd2 plays a prominent role in methylation-dependent transcriptional repression in epithelial cells. Yet transcript analyses in *Mbd2*^{-/-} mouse tissues indicate only subtle changes in transcript-level gene expression [18].

Although the mechanisms leading to promoter methylation remain to be fully

elucidated, the CpG content of promoter regions seems important. A genome-wide analysis of normal human fibroblasts has shown that the CpG content is an important predictor of methylation-dependent repression [20]. The authors subdivided the promoters into three subclasses according to their CpG content. Most promoters exhibiting a high CpG content (HCP) appeared unmethylated even when inactive, whereas low-CpG promoters (LCP) could be transcribed despite the presence of a few methylated cytosines at their 5'ends [20]. The promoters of the last group, called ICP for “intermediate CpG content”, seemed to be inactive when methylated. Taken together, these data suggest that methylation-dependent repression is restricted to subclasses of promoters. They raise questions regarding the involvement of MBD proteins in this apparently selective repression.

To further explore the relationship between DNA methylation and MBD2-dependent gene silencing, we have investigated the impact of MBD2 depletion on gene expression in the normal human embryonic lung fibroblast line MRC-5. As expected from the prominent role of Mbd2 in silencing methylated templates in mouse fibroblasts, *MBD2* is highly expressed in MRC-5 cells. This model is thus adequate for identifying and characterizing genes controlled by MBD2 in normal cells, since MRC-5 cells display characteristics of untransformed cells, exhibit a diploid karyotype, and have a limited life span.

RESULTS

Efficiency and specificity of *MBD2* knockdown by RNA interference in MRC-5 cells

Recent data suggest that the function of MBD proteins correlates, to some extent, with their level of expression in the tissue considered. Neuronal cells, for instance, show high-level *MeCP2* expression, associated with DNA-methylation-dependent repression of transcriptional noise. In these cells, loss of *MeCP2* expression results in global alteration of chromatin structure and histone methylation. Glial cells, on the other hand, show low-level *MeCP2* expression, and in these cells, loss of *MeCP2* does not cause these drastic alterations [21]. We therefore first investigated the expression level of *MBD2* and the efficiency of small interfering *MBD2* RNA (*MBD2*-siRNA) in the MRC-5 cell line. The number of *MBD2* mRNA copies per μg total RNA was determined in a quantitative competitive RT-PCR assay [22].

Untreated cells exhibited abundant *MBD2* transcripts: 7.6×10^6 mRNA molecules per μg total RNA (Figure 1). This value represents an average of 192 copies per cell, assuming that one cell contains 25 pg total RNA [23]. In MRC-5 cells exposed to *MBD2*-siRNA for 48 hours, the number of copies was reduced to 40 (Figure 1). In contrast, the expression levels of *MeCP2* and *MBD1* were unaffected in *MBD2*-depleted cells as compared to control-siRNA-treated MRC-5 cells. Thus, the dramatic (80%) decrease in the *MBD2* mRNA level in response to *MBD2*-siRNA treatment does not seem to be associated with off-target effects.

MBD2 depletion was also assessed by western blot analysis with a custom antibody directed against *MBD2*. Inhibition of *MBD2* expression resulted in the almost complete disappearance of *MBD2* protein as compared to mock-treated cells (Figure 2). To confirm the specificity of the anti-*MBD2* antibodies used, western blot analysis was also performed on extracts of HeLa cells transiently transfected with *MBD2*-siRNA or with a vector

overexpressing *Mbd2* cDNA [4]. As expected, MBD2 depletion was observed in *MBD2*-knockdown HeLa cells, and positive signals were obtained from both mock-treated and *Mbd2*-cDNA-transfected HeLa cells (data not shown). These data indicate that our custom anti-MBD2 antibodies recognize the MBD2 protein specifically and efficiently. They were therefore further used in ChIP experiments.

MBD2 is thus expressed to a high level in MRC5 cells and efficiently (80%) silenced by MBD2-siRNA. This model can therefore be used to investigate how MBD2 depletion affects transcription in normal human cells.

MBD2 depletion does not induce global gene overexpression in MRC-5 cells

The impact of MBD2 depletion on the MRC-5 transcriptome was determined on Codelink Human whole genome arrays bearing ~55,000 human and 384 non-human cDNAs or expressed sequence tags (ESTs), the latter set being used as negative controls to estimate the background noise. We performed two independent experiments and each pair of samples (MBD2-siRNA and control-siRNA) was processed independently (RNA extraction, labeling, hybridization...). On the arrays, a gene was considered expressed if the corresponding signal exceeded a threshold defined as the mean signal obtained for the negative controls plus three standard deviations (SD).

MBD2 depletion did not appear to alter transcription globally. In each sample, ~45% of the transcripts represented on the arrays were found to be expressed. There was good agreement between equivalent samples (91.5% for MBD2-siRNA-treated cells and 91% for control-siRNA-treated cells).

To accurately identify genes affected by MBD2 depletion, we focused our analysis on genes showing, in both independent experiments, at least 1.3-fold higher expression in MBD2-siRNA-treated cells than in control cells. The number of genes showing upregulated

expression in response to Mbd2 depletion was 536. This result is in keeping with the low percentage of upregulated genes observed in *Mbd2*^{-/-} mouse colon tissues [24].

Identification of genes specifically repressed by MBD2

The aim of these experiments was to characterize the genes whose expression responds to MBD2 depletion in MRC-5 cells. In this context, promoter annotation is particularly important. To eliminate falsely assigned promoters, the list of upregulated genes was filtered through the RefSeq data bank. We thus obtained a shorter list of 178 well-annotated genes. This list was used for further analysis (Table S1).

Among the 178 upregulated genes, we distinguished 42 (henceforth called “depletion-activated genes”) that appeared silent in the mock-treated cells (i.e., had a signal below the above-defined expression threshold in both duplicates) but expressed upon MBD2 depletion. These genes were strong candidates for direct MBD2-mediated silencing. To test this hypothesis, 6 of them (*LIMS2*, *EGFL7*, *PRPH2*, *FAM65A*, *DNASE1*, and *IL2RB*) were tested for expression, methylation status, and MBD2 binding in direct experiments.

We used RT-qPCR to quantify the expression of these 6 genes (Figure 3) in 3 other independent batches of MBD2-siRNA-treated MRC-5 cells. These genes appeared highly repressed in control MRC5 cells, but very low levels of expression were nevertheless detected. Significant (1.7- to 3-fold) enhancement of expression was observed upon MBD2 depletion in all three independent experiments (Figure 3). As a control, we also quantified expression of the *CDH13* gene, driven by a promoter having a low CpG content. Our RT-qPCR assays confirmed the array data indicating that MBD2 depletion does not induce *CDH13* expression (data not shown).

The methylation status of the *LIMS2*, *EGFL7*, *PRPH2*, *FAM65A*, *DNASE1* and *IL2RB*

genes was investigated by bisulfite sequencing. DNA extracted from the cell line was modified and amplified by PCR with primers specifically designed to amplify bisulfite-modified DNA sequences in the promoter region. For *DNASE1*, two alternative promoters have been described, so both of the corresponding regions were analyzed.

The percentage of methylation was determined as the number of methylated CpGs divided by the total number of CpGs. The regions analyzed all appeared highly or totally methylated, indicating that the corresponding genes have the methylation profile expected for genes undergoing methylation-dependent silencing by MBD2 (Figure 4).

We then performed chromatin immunoprecipitation experiments (ChIP) to test, in at least three independent experiments, whether MBD2 binds to the methylated promoter regions of the six selected genes (Figure 5).

Strong enrichment in *LIMS2*, *EGFL7*, *PRPH2*, *FAM65A*, *DNASE1* (both promoter regions), and *IL2RB* was observed in the fraction immunoprecipitated by anti-MBD2 antibodies. As expected, a chromatin fragment included in a CpG-free region (–1000 to –700) of the *CDH13* 5' end was not immunoprecipitated by these antibodies (Figure 5). As an additional control, ChIP experiments were performed on HeLa and constitutive MBD2–knockdown HeLa cells [4]. As expected, in immunoprecipitates from HeLa cells constitutively depleted of MBD2, the MBD2 targets *BRCA1* and *hTERT* [4, 6] were no longer detected, while enrichment in these targets was observed in immunoprecipitates from wild-type HeLa cells (data not shown). Taken together, these data indicate that the six selected “depletion-activated” genes (and very likely the other genes in this group) are MBD2 targets, strongly suggesting that induction of their expression upon MBD2 depletion is a direct effect.

Molecular function of genes upregulated upon MBD2 depletion

In addition to the above-mentioned group of 42 “depletion-activated genes”, a group of 136

genes identified as upregulated in MBD2-depleted MRC5 cells appeared to be expressed also in control cells. These are henceforth called “depletion-stimulated genes”.

In order to investigate the molecular functions of genes upregulated upon MBD2 depletion, we compared the gene ontology (GO) categories of the 42 “depletion-activated” and 136 “depletion-stimulated” genes. Screening with the GOTM (Gene Ontology tree machine: <http://bioinfo.vanderbilt.edu/gotm>) revealed distinct GO categories in which one or the other group was overrepresented (Table 1). “Depletion-stimulated” genes were mainly associated with the “coenzyme binding” and “catalytic activity” categories, while “depletion-activated” genes were overrepresented in the “DNA binding” and “transcription regulator activity” categories. These data suggest that the two groups of MBD2-regulated genes are involved in different biological pathways.

Promoter analysis of genes upregulated by MBD2 depletion

Among the characteristics of genes potentially regulated by epigenetic events, the CpG content of their 5' ends seems to play an important role. We therefore analyzed the promoter regions (positions -900 to +400 from the transcription start site, TSS) of the upregulated genes, ranking them according to the criteria defined by Weber al. [20]: HCPs (high-CpG promoters) containing a 500-bp area with a CpG ratio above 0.75 and a GC content above 55%, LCPs (low-CpG promoters) that do not contain any 500-bp area with a CpG ratio above 0.48, and ICPs (intermediate-CpG promoters) that are neither HCPs nor LCPs.

Among our two groups of genes found to be upregulated upon MBD2 depletion, promoter classes were not randomly distributed. Both the “depletion-stimulated” genes (χ^2 test, $p = 2.2 \times 10^{-19}$) and the “depletion-activated” genes (χ^2 test, $p = 2.7 \times 10^{-23}$) showed a promoter-class distribution differing significantly from that observed at whole-genome level in a cell line similar to MRC5 (the normal human embryonic lung fibroblast line WI-38) [20].

To gain further insights into this matter, we compared the distribution of promoter classes in each group of upregulated genes with the distribution of methylated promoters in the genome of WI-38 cells [20]. The distribution of HCP, ICP, and LCP, among the 136 “depletion-stimulated” genes was found not to differ significantly (χ^2 test, $p = 0.22$) from the overall distribution of methylated promoters in WI-38 cells. In contrast, the 42 “depletion-activated” genes exhibited a specific distribution of promoter classes that did not match the distribution of methylated promoters in WI-38 cells (χ^2 test, $p = 10^{-4}$). In the group of “depletion-activated” genes, HCP were underrepresented (4.8% vs. 25%) and ICP overrepresented (61.9% vs. 33%) as compared to their distribution among methylated promoters throughout the WI-38 genome.

Taken together, these data indicate that promoters containing an intermediate concentration of CpG are the main targets of DNA-methylation-dependent repression mediated by MBD2, and suggest that MBD2 depletion is not sufficient to lift the silencing of genes exhibiting a hypermethylated CpG island in their promoter region.

DISCUSSION

The correlation between DNA methylation and gene silencing is now widely recognized, and methyl-DNA binding proteins seem to play a key role in this process, mediating the recruitment of histone modification complexes to methylated regions. Among the MBP, the methyl-CpG binding protein domain proteins have been particularly studied. It is noteworthy that in mouse, *Mecp2* deletion induces neurological disorders: *Mecp2*^{-/-} mice exhibit phenotypes resembling some of the symptoms observed in Rett syndrome [19]. This human pathology, caused by mutations in *MeCP2*, is one of the most common causes of mental retardation in females [25].

The protein MBD2 is another interesting MBP. In normal fibroblasts it plays a prominent role in repressing expression from methylated templates [19]. Furthermore, it can bind to methylated DNA independently of the surrounding sequences [26]. To identify MBD2-silenced promoters in normal human fibroblasts, we have used RNA interference to analyze the impact of MBD2 depletion on transcriptional activity in MRC5 cells.

Our preliminary control experiments indicated that the chosen model was appropriate for this analysis: (i) MBD2 depletion was efficient and reproducible, the number of *MBD2* transcripts being reduced 5-fold in response to treatment with MBD2-siRNA (ii) this inhibition resulted in the almost complete disappearance of the corresponding protein (iii) MBD2 depletion was not compensated by upregulation of *MeCP1* or *MBD1* expression. In addition, inhibition of *MBD2* expression affected neither cell morphology nor cell growth. These observations are in line with the absence of any specific phenotype associated with loss of *Mbd2* in mice [15].

The number of MBD2-depletion-upregulated genes identified in our transcriptomic analysis is lower than the number identified in cancer cell lines [27]. This suggests that the

gene silencing linked to aberrant methylation patterns observed in cancer cell lines could be less stringent than “normal” DNA-methylation-dependent repression. For example, *Xist* inactivation in male mice depends on multiple layers of repressive mechanisms (Mdd2, Dnmt1, polycomb proteins...), and Mbd2 depletion does not completely alleviate *Xist* repression [18].

The redundancy of epigenetic repressors may explain why MBD2 depletion alone does not result in global transcriptomic changes. Considering the relief from repression observed for in vitro methylated templates in transient transfection experiments on fibroblasts from *Mbd2*^{-/-} mice [19], it could be that non-genomic sequences more readily evade the multiple repressive mechanisms occurring at “normal” methylated genomic sequences.

The genes upregulated upon MBD2 depletion in MRC5 cells fall into two groups: genes already expressed in mock-treated control cells (“depletion-stimulated” genes) and genes that are silent (or expressed to levels below the set threshold) in these cells. These two groups were found to differ as regards both their molecular functions (GO analysis) and their CpG content. “Depletion-stimulated” genes appeared to be frequently involved in cell metabolism, while the most significant functions of the “depletion-activated” genes were linked to the transcriptional machinery. The former genes, moreover, showed a promoter-class distribution similar to the distribution of methylated promoters at whole-genome level [20]. This suggests that the MBD2 level might not be an independent parameter, and that it might not be crucial for the regulation of these genes. Nevertheless, changes in mRNA levels may have physiological consequences, since many of the “depletion-stimulated” genes are involved in cellular metabolism.

The “depletion-activated” genes, whose expression is prevented by the presence of MBD2, exhibited a specific distribution of promoter classes. Among them, only 2 out of 42 promoters are HCP, while the “stimulated” genes exhibit a percentage (25%) of HCP similar

to the percentage of methylated HCP observed at whole genome level in the normal human lung fibroblastic cell line WI-38 [20].

Taken together, these data indicate that transcription of expressed HCP-genes can be enhanced upon MBD2 depletion, but that silent HCP-genes remain silent upon removal of this methylation-dependent repressor. This suggests that in normal fibroblasts, gene silencing mediated by CpG island methylation is associated with a tighter repressive chromatin structure than the silencing of genes having ICP or LCP promoters. In cancer cell lines, a higher proportion of genes, exhibiting aberrantly methylated CpG islands, seems to be “stimulated” upon MBD2 depletion. This suggests the possibility that abnormal methylation may lead to weaker repression, but this hypothesis requires verification, through a detailed comparison of the chromatin structures of “aberrantly methylated” and “normally methylated” CpG islands.

MATERIALS AND METHODS

Cell culture

MRC-5 cells (a normal human embryonic lung fibroblast line) and HeLa cells (a cervical adenocarcinoma cell line), were obtained from the American Type Culture Collection (ATCC, Rockville, MD.) and grown in DMEM supplemented with 10% heat-inactivated fetal bovine serum (Lonza, Vervier, Belgium) at 37°C in a humidified 5% CO₂ atmosphere.

Transient transfection

Transient transfections of HeLa cells with MBD2-siRNA duplexes, control (scrambled) siRNA, or the pRev-Mbd2 expression vector were carried out as previously described [6]. Cells were harvested 48 hours after transfection.

MRC-5 cells were transiently transfected with 30 nM MBD2-siRNA duplexes or control siRNA by means of LipofectamineTM RNAiMAX (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Briefly, 800,000 cells were plated in a dish (100 mm in diameter) in a medium without antibiotics 24 hours before transfection. Then dilutions of siRNA duplexes and lipofectamine complexes in OptiMEM medium (Invitrogen) were mixed and incubated for 15 minutes before their addition to cell culture media. Cells were harvested 48 hours after transfection.

The efficiency of siRNA treatments was determined from the *MBD2* mRNA and MBD2 protein levels. Total RNA was extracted with the RNeasy Mini Kit (Qiagen, Courtaboeuf, France), and quantified by densitometry [6]. *MBD2* transcripts were quantified by quantitative competitive RT-PCR as previously described [22]. MBD2 protein levels were analyzed by western blotting. Briefly, after isolation, nuclei were lysed in a buffer containing

50 mM Tris, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.1% SDS, and protease inhibitors. Nuclear proteins (50 µg) were separated by electrophoresis on 12% polyacrylamide gels [22]. The MBD2 protein was detected with purified custom rabbit polyclonal antibodies diluted in TBS + 0.05% Tween + 2% BSA. Custom antibodies were obtained after immunization with peptides corresponding to the N-terminal part of the MBD2 protein (Covalab, Villeurbanne, France). The secondary antibodies used were polyclonal goat anti-rabbit immunoglobulins/HR, (Dakocytomation, Trappes, France) diluted 1:3000 in TBS + 0.05% Tween + 2% BSA. As a loading control the same membrane was probed with an antibody directed against the human Ku80 protein (Purified Mouse Anti-Human Ku80, BD Biosciences, Le Pont de Claix, France). Immune complexes were detected with the ECL system (Amersham, Saclay, France).

Reverse-transcription-PCR analysis

Total RNA was extracted from the cell lines with the RNeasy Mini Kit (Qiagen). After extraction, the integrity of the total RNA was examined on a 1.2% agarose gel containing 1 µg/ml ethidium bromide. To quantify RNA quantified by spectrofluorimetry, the VersaFluor™ Fluorometer (Biorad, Ivry, France) and the RiboGreen reagent (Molecular Probes, Interchim, Montluçon, France) were used. After cDNA synthesis (iScript cDNA synthesis, Bio-Rad, Marnes-la-Coquette, France), Real-time RT-PCR was carried out to quantify the gene transcripts, with iQ SYBR Green Supermix on a MiniOpticon system according to the manufacturer's instructions (Bio-Rad). *PBGD* mRNA was used as a reference. The primer sequences used for reverse-transcription PCR are available in Supplemental Table S2.

Genome-wide expression analysis

Genome-wide gene expression patterns of MBD2-depleted and control cells were analyzed on CodeLink™ human whole genome bioarrays (Amersham) representing about 55,000 transcripts. After synthesis of cDNAs corresponding to the polyA+ RNA population, cRNAs were produced by in vitro transcription, labeled, and hybridized to arrays by ProfileXpert service (ProfileExpert, Bron, France) according to Amersham's instructions.

Analysis of CpG and G+C contents of promoter regions

In order to compare our own data with published data [20], we analyzed the CpG and G+C contents of the genes' 5' regions, from bp -900 to bp +400 relative to the transcriptional start site. CpGProd (<http://pbil.univ-lyon1.fr/software/cpgprod.html>) was used to classify the genes as high-CpG promoters (HCP), low-CpG promoters (LCP), or intermediate-CpG promoters (ICP) according to their GC content and to the ratio of observed versus expected CpG dinucleotides.

DNA methylation analysis

Sodium bisulfite modification was performed as described previously [28]. The primers used for PCR amplification of the promoter regions of the *PRPH2*, *EGFL7*, *IL2RB*, *DNASE1* (distal and proximal promoters), *LIMS2*, and *FAM65A* genes are shown in supplemental Table S1. After purification, converted DNAs were amplified by PCR. The heating/cooling program was 15 min at 95°C for activation of HotStar *Taq* DNA polymerase (Qiagen, Courtaboeuf, France), 40 cycles (30 s at 94°C for denaturation, 1 min at an appropriate temperature for annealing, and 1 min 30 s at 72°C for extension), with a final extension step (10 min at 72°C). Amplified products were analyzed by automatic sequencing (Biofidal, Lyon, France) with PCR primers.

Chromatin immunoprecipitation assays

Chromatin immunoprecipitation assays were carried out as previously described [4, 6]. Briefly, cross-linked chromatin fragments 300 to 600 bp long were precipitated with antibodies directed against MBD2. Purified DNAs from the input and bound fractions were quantified, and equal amounts of each fraction were amplified by qPCR. Enrichment was determined from the ratio of bound versus input for each fragment.

ACKNOWLEDGMENTS

L. Perriaud is the recipient of a fellowship from Association pour la Recherche sur le Cancer. The present work was supported by the Ligue Nationale contre le Cancer (Comité du Rhône).

REFERENCES

1. Klose RJ, Bird AP (2006) Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* 31: 89-97.
2. Sasai N, Nakao M, Defossez PA (2010) Sequence-specific recognition of methylated DNA by human zinc-finger proteins. *Nucleic Acids Res.*
3. Bird AP, Wolffe AP (1999) Methylation-induced repression--belts, braces, and chromatin. *Cell* 99: 451-454.
4. Auriol E, Billard LM, Magdinier F, Dante R (2005) Specific binding of the methyl binding domain protein 2 at the BRCA1-NBR2 locus. *Nucleic Acids Res* 33: 4243-4254.
5. Ballestar E, Paz MF, Valle L, Wei S, Fraga MF *et al.* (2003) Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer. *Embo J* 22: 6335-6345.
6. Chatagnon A, Bougel S, Perriaud L, Lachuer J, Benhattar J *et al.* (2009) Specific association between the methyl-CpG-binding domain protein 2 and the hypermethylated region of the human telomerase reverse transcriptase promoter in cancer cells. *Carcinogenesis* 30: 28-34.
7. Koizume S, Tachibana K, Sekiya T, Hirohashi S, Shiraishi M (2002) Heterogeneity in the modification and involvement of chromatin components of the CpG island of the silenced human CDH1 gene in cancer cells. *Nucleic Acids Res* 30: 4770-4780.
8. Kondo E, Gu Z, Horii A, Fukushige S (2005) The thymine DNA glycosylase MBD4 represses transcription and is associated with methylated p16(INK4a) and hMLH1 genes. *Mol Cell Biol* 25: 4388-4396.

9. Lin X, Asgari K, Putzi MJ, Gage WR, Yu X *et al.* (2001) Reversal of GSTP1 CpG island hypermethylation and reactivation of pi-class glutathione S-transferase (GSTP1) expression in human prostate cancer cells by treatment with procainamide. *Cancer Res* 61: 8611-8616.
10. Lopez-Serra L, Ballestar E, Fraga MF, Alaminos M, Setien F *et al.* (2006) A profile of methyl-CpG binding domain protein occupancy of hypermethylated promoter CpG islands of tumor suppressor genes in human cancer. *Cancer Res* 66: 8342-8346.
11. Magdinier F, Wolffe AP (2001) Selective association of the methyl-CpG binding protein MBD2 with the silent p14/p16 locus in human neoplasia. *Proc Natl Acad Sci U S A* 98: 4990-4995.
12. Nguyen CT, Gonzales FA, Jones PA (2001) Altered chromatin structure associated with methylation-induced gene silencing in cancer cells: correlation of accessibility, methylation, MeCP2 binding and acetylation. *Nucleic Acids Res* 29: 4598-4606.
13. Macdonald JL, Verster A, Berndt A, Roskams AJ (2010) MBD2 and MeCP2 regulate distinct transitions in the stage-specific differentiation of olfactory receptor neurons. *Mol Cell Neurosci* 44: 55-67.
14. Martin Caballero I, Hansen J, Leaford D, Pollard S, Hendrich BD (2009) The methyl-CpG binding proteins Mecp2, Mbd2 and Kaiso are dispensable for mouse embryogenesis, but play a redundant function in neural differentiation. *PLoS One* 4: e4315.
15. Hendrich B, Guy J, Ramsahoye B, Wilson VA, Bird A (2001) Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. *Genes Dev* 15: 710-723.
16. Hutchins AS, Mullen AC, Lee HW, Sykes KJ, High FA *et al.* (2002) Gene silencing quantitatively controls the function of a developmental trans-activator. *Mol Cell* 10:

- 81-91.
17. Berger J, Sansom O, Clarke A, Bird A (2007) MBD2 is required for correct spatial gene expression in the gut. *Mol Cell Biol* 27: 4049-4057.
 18. Barr H, Hermann A, Berger J, Tsai HH, Adie K *et al.* (2007) Mbd2 contributes to DNA methylation-directed repression of the Xist gene. *Mol Cell Biol* 27: 3750-3757.
 19. Guy J, Hendrich B, Holmes M, Martin JE, Bird A (2001) A mouse *Mecp2*-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat Genet* 27: 322-326.
 20. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S *et al.* (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* 39: 457-466.
 21. Skene PJ, Illingworth RS, Webb S, Kerr AR, James KD *et al.* (2010) Neuronal MeCP2 Is Expressed at Near Histone-Octamer Levels and Globally Alters the Chromatin State. *Mol Cell* 37: 457-468.
 22. Billard LM, Magdinier F, Lenoir GM, Frappart L, Dante R (2002) MeCP2 and MBD2 expression during normal and pathological growth of the human mammary gland. *Oncogene* 21: 2704-2712.
 23. Wilson CA, Payton MN, Elliott GS, Buaas FW, Cajulis EE *et al.* (1997) Differential subcellular localization, expression and biological toxicity of BRCA1 and the splice variant BRCA1-delta11b. *Oncogene* 14: 1-16.
 24. Sansom OJ, Berger J, Bishop SM, Hendrich B, Bird A *et al.* (2003) Deficiency of Mbd2 suppresses intestinal tumorigenesis. *Nat Genet* 34: 145-147.
 25. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U *et al.* (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* 23: 185-188.

26. Klose RJ, Sarraf SA, Schmiedeberg L, McDermott SM, Stancheva I *et al.* (2005) DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. *Mol Cell* 19: 667-678.
27. Lopez-Serra L, Ballestar E, Ropero S, Setien F, Billard LM *et al.* (2008) Unmasking of epigenetically silenced candidate tumor suppressor genes by removal of methyl-CpG-binding domain proteins. *Oncogene* 27: 3556-3566.
28. Magdinier F, Billard LM, Wittmann G, Frappart L, Benchaib M *et al.* (2000) Regional methylation of the 5' end CpG island of BRCA1 is associated with reduced gene expression in human somatic cells. *Faseb J* 14: 1585-1594.

Figure legends

Figure 1. *MBD2* mRNA and protein quantification in *MBD2*-siRNA-treated MRC5 cells.

(A) MRC5 cells were transiently transfected with either *MBD2*-siRNA or a negative control (scrambled) siRNA. 48 hours after transfection, RNA was extracted and the efficiency of the *MBD2*-siRNA treatment was determined by quantification of *MBD2* mRNA as previously described [22]. The number of *MBD2* transcripts was determined from three independent experiments. (B) Immunoblot analysis of *MBD2* proteins in MRC5 cells treated with *MBD2*-siRNA or (scrambled) control siRNA. The blots were probed with a custom rabbit polyclonal antibody against *MBD2*. Then the same membrane was probed with purified mouse anti-human Ku80 as a loading control. Arrows indicate the two isoforms of *MBD2*, *MBD2a* (55 kDa) and *MBD2b* (32 kDa).

Figure 2. Transcriptional activation in *MBD2*-depleted MRC5 cells.

Among the genes whose expression was activated upon *MBD2* depletion, identified from chip analysis, six of them, *LIMS2*, *EGFL7*, *PRPH2*, *FAM65A*, *DNASE1*, and *IL2RB*, were analyzed by RT-Q-PCR. Fold change in gene expression was determined from the ratio between *MBD2*-depleted cells (*MBD2*-siRNA-transfected cells) and control cells (scrambled-siRNA-transfected cells). Gene expression was monitored by quantitative RT-PCR 48 h after transfection. Mean values (\pm SEM) obtained from at least three independent transfection experiments are shown.

Figure 3. DNA methylation patterns of *LIMS2*, *EGFL7*, *PRPH2*, *FAM65A*, *DNASE1*, and *IL2RB* in MRC5 cells.

(A) Bisulfite sequencing of 6 genes upregulated upon *MBD2* depletion in MRC5 cells. Bisulfite-treated DNA was PCR amplified and PCR products were then directly sequenced.

For *DNASE1*, two promoter regions were analyzed: P, the proximal promoter and D, the distal promoter. The percentage of methylation was determined by dividing the number of methylated CpGs by the number of CpGs in the corresponding unmodified PCR fragment.

(B) Number of methylated CpGs in PCR products. The lengths of the PCR products were: 268 bp, *LIMS2*; 199 bp, *EGFL7*; 289 bp, *PRPH2*; 288 bp, *FAM65A*; 236 bp *DNASE1 P*, 264 bp *DNASE1 D*; and 267 bp, *IL2RB*.

Figure 4. MBD2-associated genes upregulated upon MBD2 depletion.

MBD2 binding to the promoter region of the six genes *LIMS2*, *EGFL7*, *PRPH2*, *FAM65A*, *DNASE1*, and *IL2RB*, characterized as regards their expression and methylation status upon MBD2 depletion. As a control, a promoter region of *CDH13* was also analyzed: *CDH13* expression was unaffected by MBD2 depletion, and the promoter region analyzed did contain CpG sequences. Cross-linked chromatin was immunoprecipitated with rabbit polyclonal anti-MBD2 antibodies. After immunoprecipitation, purified DNA from the input and bound fractions was quantified and 0.5 ng of this DNA was amplified by qPCR analysis (error bars, SEM from at least 3 independent experiments). The bound/input ratio is shown for each gene analyzed. Enrichment in the bound fractions was observed for *LIMS2*, *EGFL7*, *PRPH2*, *FAM65A*, *DNASE1*, and *IL2RB*, and as expected, depletion in the bound fraction was observed for *CDH13*.

Figure 5. Classification of promoters by their CpG and G+C contents.

Genes upregulated upon MBD2 depletion were classified according to the CpG and G+C contents of their 5' end regions (positions -900 to +400 relative to their TSS). Three classes were defined: HCP contains a 500-bp area with a CpG ratio above 0.75 and a GC content above 55%; LCP does not contain any 500-bp area with a CpG ratio above 0.48; ICP is

neither HCP nor LCP [20]. Pie charts show the distribution of the 3 promoter classes among the genes upregulated upon MBD2 depletion in MRC5 cells. **(A)** upregulated genes, i.e. “depletion-stimulated” genes (expressed in control cells) + “depletion-activated” genes (no transcripts detected in control cells); **(B)** “depletion-stimulated” genes; **(C)** “depletion-activated” genes. In all three groups (A, B, and C), the distribution of promoter classes was significantly different from the distribution observed in the whole genome, χ^2 test, $p = 1.2 \times 10^{-39}$, $p = 2.2 \times 10^{-19}$, and $p = 1.9 \times 10^{-23}$, respectively. When promoter-class distributions in groups A, B, and C were compared with the distribution of classes among the methylated promoters observed in the normal human embryonic fibroblast line WI-38, groups A and B were not significantly different from WI-38, χ^2 test, $p = 0.09$, $p = 0.22$, respectively, while group C exhibited a specific promoter distribution, χ^2 test, $p = 1.7 \times 10^{-3}$.

“Depletion-stimulated” genes (136)		
Gene Ontology category	p-value (GOTM)	genes in whole dataset
coenzyme binding (GO:0050662)	0.0007	7/136
catalytic activity (GO:0003824)	0.0089	51/136
chromatin binding (GO:0003682)	0.0048	5/136
calcium ion binding (GO:0005509)	0.005	15/136
cytoskeletal protein binding (GO:0008092)	0.0118	9/136
“Depletion-activated » genes (42)		
DNA binding (GO:0003677)	0.004	12/42
transcription regulator activity (GO:0030528)	0.0161	8/42
RNA polymerase II transcription factor activity (GO:0003702)	0.0153	3/42

TABLE 1. Top molecular functions of genes upregulated upon MBD2 depletion in HeLa cells.

The upregulated genes were analyzed with the GOTM (Gene Ontology tree machine:

<http://bioinfo.vanderbilt.edu/gotm>).

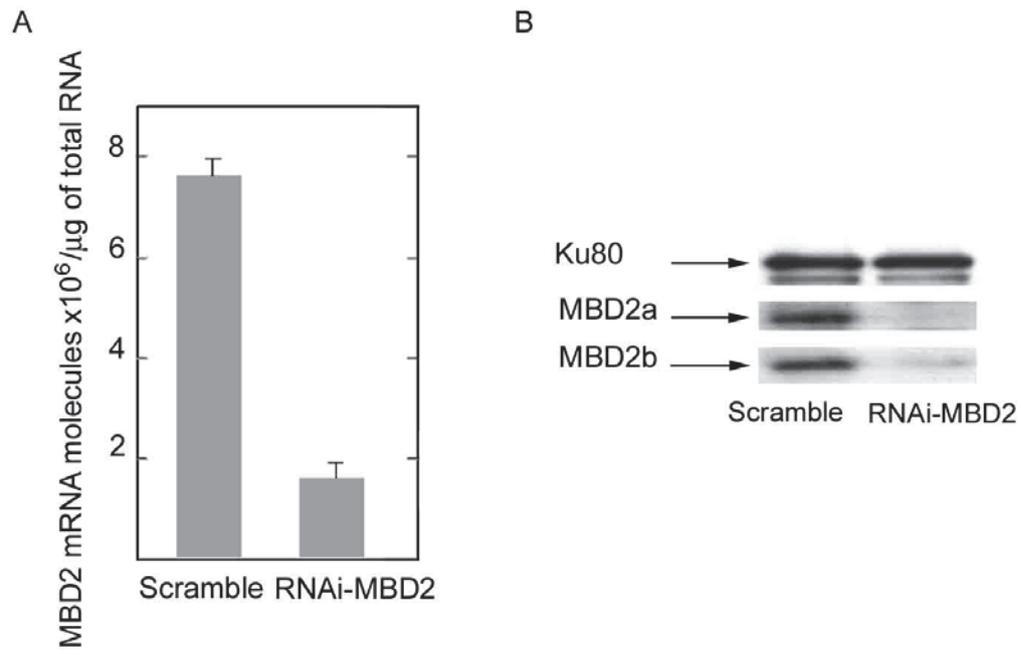


Figure 1

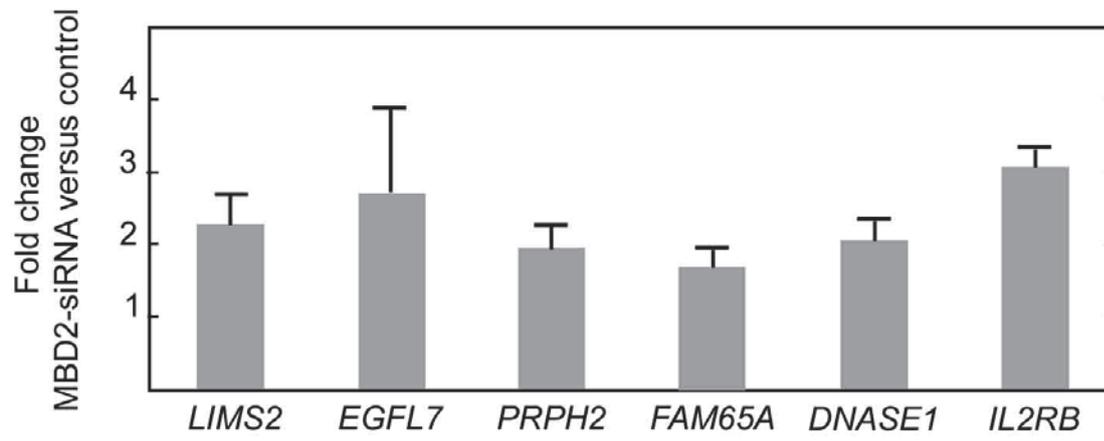


Figure 2

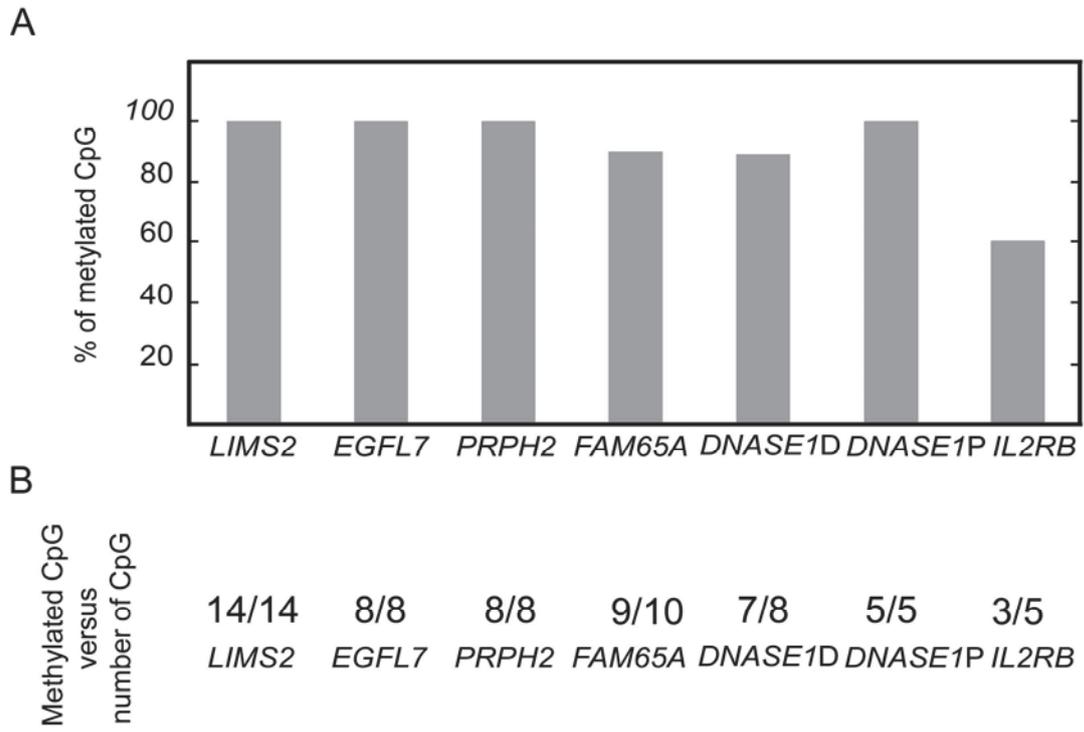


Figure 3

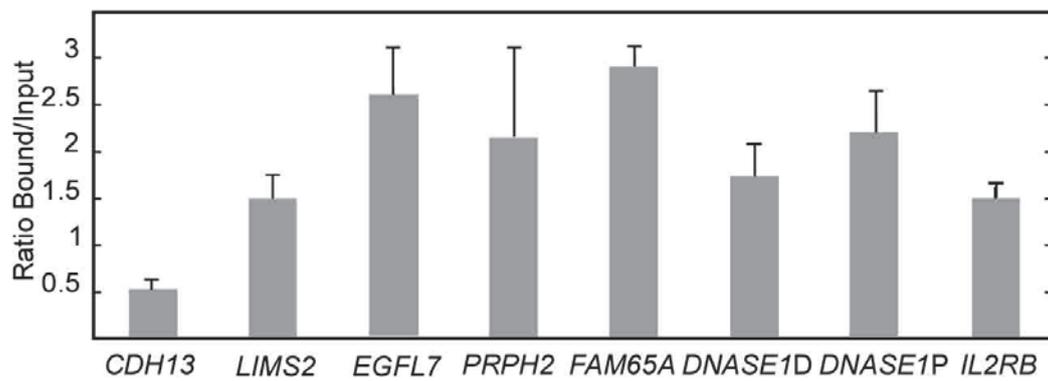


Figure 4

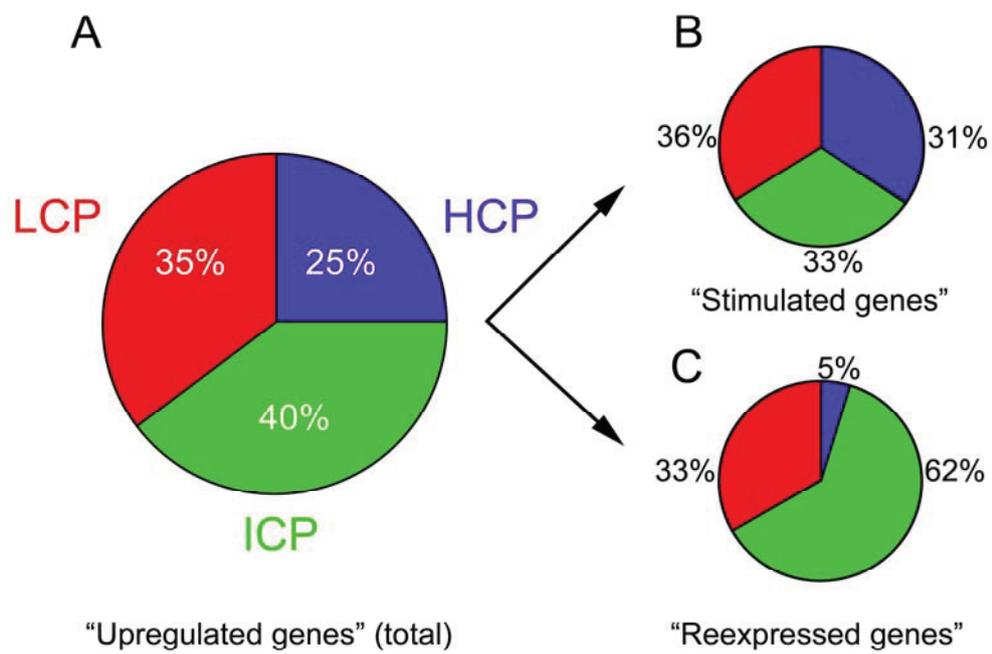


Figure 5

Article 3

Methyl-CpG Binding Domain Protein 2, a DNA Methylation- Dependent Gene Silencer, Genome-Wide Analysis

Amandine Chatagnon, **Laury Perriaud**, Nicolas Nazaret, Séverine Croze, Jean Benhattar,
Joël Lachuer, and Robert Dante

A soumettre de nouveau (prévu à *Nucleic Acids Research*)

Contexte de l'étude

De nombreux gènes réprimés par la présence de la protéine MBD2 sur leurs promoteurs méthylés ont déjà été identifiés par des approches « gène candidat » dans diverses lignées cellulaires cancéreuses, notamment dans les cellules humaines du cancer du col de l'utérus HeLa, et tissus tumoraux. De plus, la protéine MBD2 semble, au moins pour certains cancers, impliquée dans le processus tumoral. Aussi, il paraît nécessaire d'**identifier les cibles spécifiques de la protéine MBD2 à l'échelle du génome au sein de lignées cellulaires et/ou tissus cancéreux.**

En collaboration avec l'équipe du Dr Esteller, l'équipe a déjà mis en évidence à l'échelle du génome une relation directe entre les protéines à MBD MBD1, MBD2 et MeCP2 et la répression transcriptionnelle à partir des cellules HeLa (Lopez-Serra, Ballestar et al. 2008). La protéine MBD2 semble être un répresseur majeur car responsable à 93% de la ré-expression génique observée suite à la triple invalidation des ARNm *MBD1*, *MBD2* et *MeCP2*. Les gènes touchés par l'absence de MBD2 (~900) présentent en majorité un îlot de CpG (~75%) à leur extrémité 5' et une faible proportion appartient à des gènes soumis à l'empreinte parentale et à l'inactivation du chromosome X.

De plus, la protéine MBD2 a été largement décrite comme répresseur dépendant de la méthylation pour sa fixation. Cependant, certaines études récentes suggèrent que MBD2 peut aussi jouer le rôle d'activateur transcriptionnel lorsqu'elle est recrutée par le complexe CREB via RHA (RNA helicase A) (Fujita, Fujii et al. 2003) ou qu'elle recrute TACC3/pCAF (Angrisano, Lembo et al. 2006) ou MBD-in (Lembo, Pero et al. 2003) sur des sites méthylés. Une étude récente dans les neurones a même montré que la majorité des sites liés par la protéine MeCP2 étaient actifs alors que jusqu'à présent elle ne paraissait être qu'un répresseur (Yasui, Peddada et al. 2007).

Stratégie de l'étude

Il paraît **incontournable d'élargir l'analyse des profils de fixation de la protéine MBD2 et des profils de méthylation de l'ADN à l'ensemble des promoteurs du génome d'une part pour identifier les cibles spécifiques de MBD2 et d'autre part pour caractériser son rôle.**

Dans ce but, nous avons choisi d'utiliser les **cellules HeLa comme modèle** et de **développer des approches globales directes de « Methylation DNA Precipitation-on-chip »** afin d'identifier les sites présentant une méthylation et de **« Chromatin ImmunoPrecipitation-on-chip »** pour cartographier les sites liés par MBD2 et les sites où l'ARN polymérase II est absente, marquant une absence de transcription.

Originalité des approches

L'approche « ChIP-on-chip » constitue une **technique récente permettant de déterminer à l'échelle globale du génome les sites de l'ADN directement liés par différentes protéines**. Comme la protéine MBD2 doit jouer un rôle dans l'initiation de la répression et/ou l'activation transcriptionnelle suite à sa fixation sur l'ADN, nous avons choisi d'analyser les ADN issus des immunoprécipitations de chromatine (ChIP) avec des anticorps dirigés soit **contre la protéine MBD2** soit **contre l'ARN polymérase II** par hybridation sur des **puces « Affymetrix Human promoter 1.0R Arrays »** permettant l'étude de **25.500 promoteurs**. Chaque promoteur est couvert par des sondes de 25pb espacées d'environ 10pb sur -7,45kb à +2,5kb par rapport au site d'initiation de la transcription et pour 1.300 promoteurs de gènes associés au cancer, 1,3kb sont ajoutés en amont des -7,45kb.

De plus, jusqu'à présent, les principaux outils utilisés pour sélectionner l'ADN méthylé avant hybridation sur puces à ADN étaient les anticorps anti-5-méthylcytosine nécessitant la dénaturation de l'ADN et ciblant préférentiellement des régions richement méthylées. Pour essayer de s'affranchir de ces inconvénients, **j'ai développé une technique utilisant une protéine recombinante nommée 4xMBD (Figure...)** construite par Jørgensen et ses collaborateurs (Jorgensen, Adie et al. 2006), **reconnaissant les CpG méthylés dans l'ADN double brin *in vitro* et *in vivo*** via ces 4 domaines MBD de MBD1. En effet, cette protéine a la particularité de reconnaître des sondes contenant au moins un CpG méthylé.

Les « **puces promoteurs** » ont été hybridées avec des ADN issus (i) **d'une nouvelle technique d'immunoprécipitation d'ADN (IP) à partir d'une protéine chimérique recombinante (vecteur fourni par A. Bird) pour étudier les sites de méthylation** et (ii) **d'une technique d'immunoprécipitation de chromatine (ChIP) avec des anticorps dirigés contre la protéine MBD2 pour étudier ses cibles directes ou contre l'ARN polymérase II pour identifier les régions sans activité transcriptionnelle.**

Résultats

Mise en place des « ChIP-on-chip » et « MeDP-on-chip »

Après une mise au point de la production de la protéine 4xMBD, **j'ai vérifié sa fonctionnalité et sa spécificité de liaison à l'ADN méthylé**, par retard sur gel, avec des sondes de même séquence biotinylées contenant 5 CpG méthylés ou non et utiliser une protéine dérivée de 4xMBD ayant une mutation R22A dans chaque domaine MBD comme contrôle négatif. La protéine 4xMBD produite apparaît spécifique des sites méthylés *in vitro*. Puis **j'ai mis au point une précipitation d'ADN de cellules HeLa avec cette protéine pour sélectionner l'ADN méthylé**. Cette technique consiste à fragmenter l'ADN par sonication, à l'incuber avec la protéine recombinante 4xMBD avant de récupérer les complexes avec des billes magnétiques d'agarose recouvertes de groupements nickels accrochant l'étiquette histidine des protéines 4xMBD. Puis ces complexes sont digérés par la protéinase K avant que l'ADN précipité ne soit purifié.

La spécificité de ces IP a été évaluée par comparaison de l'enrichissement en ADN méthylé des fractions précipitées ou « bound » par rapport aux fractions initiales ou « input » en analysant par PCRq des régions dont l'état de méthylation, de présence/absence de MBD2 et l'état transcriptionnel avaient été déterminé auparavant. Ainsi, pour des régions méthylées pauvre en CpG (promoteur de *pS2*) et riche en CpG (îlot de CpG sur le promoteur de *BRCA1*), j'ai obtenu des enrichissements alors que pour des régions non méthylées (exon1 de *BRCA1* et îlot de CpG sur le promoteur de *RASSF1A*), j'ai obtenu des appauvrissements. Ces résultats **montrent que l'ADN méthylé est bien spécifiquement précipité par la protéine 4xMBD *in vivo*** et que **la fixation de la protéine 4xMBD s'effectue sans biais dû à la densité en CpG**.

Après ces validations, les ADN issus des fractions précipitées et initiales ont été confiés à la plateforme ProfileXpert (Lyon) afin qu'ils soient amplifiés par PCR aléatoire, marqués puis hybridés sur des puces promoteurs « Affymetrix Human promoter 1.0R Arrays ». En collaboration avec la plateforme, nous avons ensuite effectué l'analyse des données grâce aux logiciels GeneChip® Operating Software (GCOS), Affymetrix® Tiling Analysis Software (TAS), Integrated Genome Browser (IGB) et CisGenome. En effet, GCOS permet de collecter les intensités de signaux et de les situer sur le génome. TAS sert à

normaliser les signaux (bound/input). IGB permet de visualiser les profils de gènes et de fixer un seuil à partir des régions étudiées précédemment. Ce seuil servira sous TAS à définir toutes les zones appauvries ou enrichies en méthylation. Enfin, CisGenome permet d'annoter les régions afin de les associer aux gènes. Ces données ont enfin été filtrées (élimination de doublons, des données issues des chromosomes sexuels, ...) à l'aide de logiciels Microsoft Excel et Access.

Une seconde validation du seuil a été faite avec les données de la littérature et par la comparaison avec des gènes analysés pour leur méthylation par une méthode indépendante développée par l'équipe du Dr Benhattar (Lausanne, Suisse) consistant à digérer l'ADN avec l'enzyme sensible à la méthylation *CfoI* avant d'amplifier simultanément des fragments par une PCR multiple (MLPA).

En parallèle, des CHIP ont été **réalisées avec des anticorps dirigés contre la protéine MBD2 ou contre l'ARN polymérase II**. La stratégie de validation et de fixation des seuils des puces reste la même que celle employée pour les précipitations d'ADN méthylé.

Les résultats des puces concernant la présence de la protéine MBD2 et l'absence de l'ARN polymérase II sont concordants à ~78% respectivement avec les données issues d'autres CHIP effectuées par l'équipe du Dr Esteller et les données issues de puces d'expression CodeLink effectuées par nos soins de la même façon que décrite pour les cellules MRC5 (article 1).

Analyses et croisements des données issues des puces

Les résultats issus de « CHIP anti-MBD2 -on-chip » permettent d'**identifier les cibles spécifiques de la protéine MBD2**. Ils révèlent que **~36% des promoteurs autosomaux de la puce présentent au moins un site fixé par la protéine MBD2**. Si ces sites peuvent occuper n'importe quelle position sur la région promotrice couverte par les sondes, l'étude de la distribution de la protéine MBD2 sur les promoteurs du chromosome 1 (chromosome contenant le plus de gènes) montre que la **protéine MBD2 se fixe préférentiellement sur la région comprise entre -700 et +200pb autour du site d'initiation de la transcription**. Ceci concorde avec un rôle dans l'initiation de la transcription. Cette région a ensuite été retenue pour les autres analyses.

Comme la protéine MBD2 est surtout décrite pour son rôle dans la répression dépendante de la méthylation, une analyse des sites méthylés et réprimés a ensuite été réalisée. La distribution des sites méthylés, leur taille et leur densité en CpG, ont permis d'établir la **définition d'une « région méthylée »: il s'agit une région d'au moins 200pb contenant au minimum 8 CpG méthylés et localisée dans l'intervalle -700 à +200pb autour du TSS**. De plus, si la **distribution de la méthylation semble moins importante dans la région autour du TSS**, il apparaît néanmoins une **corrélation statistiquement significative entre la méthylation et la présence de la protéine MBD2** (test χ^2 : $p = 1.1 \times 10^{-145}$), suggérant un **recrutement spécifique de la protéine MBD2 à ces sites**. Nous avons donc cherché à mieux caractériser les sites méthylés fixés par cette protéine.

Il semble que **la présence d'îlot de CpG ne présente pas un facteur de préférence et aucun motif de séquence consensus n'a été identifié**, si ce n'est un motif CG voire peut être CCG. D'autre part, le croisement des données issues des deux « ChIP-on-chip » révèle que **la majorité des sites positifs pour MBD2 sont associés à une absence de l'ARN polymérase**, suggérant que **MBD2 est impliquée dans la répression transcriptionnelle de ces gènes**.

Concernant le recoupement des données issues des « ChIP anti-ARN polymérase II - on-chip » et « MeDP-on-chip », il apparaît que près de 70% des sites silencieux sont méthylés et inversement ce qui suggère une **corrélation inverse entre la présence de méthylation et l'expression génique et montre la forte influence de la méthylation dans les mécanismes répressifs**.

D'autre part, la comparaison des données issues de chaque paramètre met en évidence que **51% des sites fixés par MBD2 sont méthylés et silencieux**, ce qui représente 18% des promoteurs totaux de la puce. De plus, **~41% des gènes méthylés et silencieux sont fixés par la protéine MBD2**. Tous ces résultats donnent à cette protéine un **important rôle de répresseur dépendant de la méthylation dans ces cellules**.

Pour évaluer la distance de répression initiée par la liaison de MBD2 à des sites méthylés, nous avons étudié la distribution des sites méthylés, liés par MBD2 et réprimés en fonction de la distance par rapport au TSS. Il apparaît que **la liaison de MBD2 autour du TSS a un rôle majeur dans la répression transcriptionnelle** et probablement plutôt dans son initiation.

Enfin, l'étude de la fonctionnalité moléculaire et cellulaire par le logiciel Ingenuity Pathway Analysis des protéines codées par les **gènes réprimés via la présence de MBD2 sur des sites méthylés semble indiquer qu'elles sont impliquées dans l'expression génique, le cycle cellulaire et le contrôle de la mort cellulaire.**

Ainsi, au cours de cette étude, j'ai donc **mis au point et validé une technique de précipitation d'ADN.** Nous avons également **développé dans l'équipe des techniques « MeDP-on-chip » et « ChIP-on-chip »,** vérifié qu'elles fonctionnaient et **mis en place une stratégie bioinformatique pour analyser** les « puces promoteurs » et **croiser les données** afin d'identifier les relations existantes entre la fixation de la protéine MBD2, la méthylation de l'ADN et l'inactivité transcriptionnelle. Dans les cellules HeLa, la protéine MBD2 parait donc être un répresseur méthyl-dépendant véritable et majeur dans ce mécanisme.

Ce type d'étude a été appliqué à d'autres modèles/lignées cellulaires.

**Methyl-CpG Binding Domain Protein 2, a DNA Methylation-Dependent Gene Silencer,
Genome-Wide Analysis**

Amandine Chatagnon¹, Laury Perriaud¹, Nicolas Nazaret², Séverine Croze², Jean Benhattar³,
Joël Lachuer², and Robert Dante¹

¹ INSERM, U590, Lyon, F-69008, France;

² ProfileXpert, INSERM U842, Bron, F-69676, France.

³ Institut de Pathologie, Centre Hospitalier, Universitaire Vaudois and University of
Lausanne, Lausanne, Switzerland ;

Correspondence to: Dr R. Dante, INSERM U590, Oncogénèse et Progression Tumorale,
Centre Léon Bérard, 28 rue Laënnec, 69373 Lyon Cedex 08, FRANCE.

E-mail: dante@univ-lyon1.fr

Key words: MBD2/ DNA methylation/ transcriptional regulation/ ChIP and MeDP on chip

Reviewer's access to ArrayExpress :
Username: Reviewer_E-MEXP-2634
Password: 1269266717622

ABSTRACT

Methyl-CpG-Binding Domain (MBD) proteins are thought to be key molecules in the interpretation of DNA methylation signals that lead to gene silencing through the recruitment of chromatin remodeling complexes. In cancer, one member of this family, MBD2, seems to play an important role in the repression of hypermethylated genes. However, recent data suggest an antithetic role for this protein. To elucidate, at whole genome level, the functional relevance of MBD2 in transcriptional regulation, we analyzed, in HeLa cells, MBD2 binding sites along with DNA methylation patterns and RNA polymerase II occupancy on array representing 25,500 promoter regions. This first whole genome mapping indicated that MBD2 associated methylated regions ($p = 1.1 \times 10^{-145}$). Transcriptional silencing was also correlated with DNA methylation ($p = 7.9 \times 10^{-212}$) and a non-random association ($p < 0.0001$) between silent genes, DNA methylation and MBD2 binding was observed. Furthermore, an increase in the distance between MBD2 binding sites and the transcription start sites lessened the strength of this association, suggesting that MBD2 may repress transcriptional initiation. Among the silent genes, approximately 42% exhibited methylated sites bound by MBD2, indicating that MBD2 may be predominant in the repression associated with DNA methylation, in HeLa cells.

INTRODUCTION

Since the discovery of an inverse correlation between the methylation of the cytosine in 5'-C-G-3' dinucleotide (also called CpG) and transcriptional activity, the potential relationship between DNA methylation and gene silencing has been investigated in many organisms (1). A large part of these studies has been performed in human and mouse cell lines and has shown that specific proteins or complexes can recognize methylated DNA and mediate transcriptional repression by changing the local chromatin structure (2).

The first protein able to associate methylated CpGs, independently of DNA sequences, was identified over 15 years ago and called MeCP2 (3). Further studies have shown that MeCP2 belongs to a protein family of 5 members, including MBD1, MBD2, MBD3, and MBD4. These proteins share, with the exception of MBD3, a same functional methyl-CpG-binding domain (MBD) responsible of their binding specificity (4). Three of them, MBD1, MBD2, and MeCP2, mediate transcriptional repression by targeting chromatin remodeling co-repressors complexes to methylated DNA sequences, thereby establishing a repressive chromatin state (5-7). Although MBD3 lacks methyl-CpG-binding activity due to amino acid changes at critical positions (4), this protein is part of the histone deacetylase (HDAC) and chromatin remodeling Mi2/NuRD complex (6, 8). The last member of the MBD protein family, MBD4, was also found to be involved in gene silencing (9, 10). However, MBD4 was first identified as thymine glycosylase and seems to be primarily involved in DNA repair (11).

The MBD proteins have been thus established to be critical mediators of many biological processes such as X-chromosome inactivation (12) and genomic imprinting (13). Surprisingly, *Mbd1*, *Mbd2*, and *Mecp2* knockout mouse experiments indicated that the loss of these genes is not associated with a global mis-expression of methylated genes and mouse do not exhibit dramatic phenotypes (14-16). However, disruptions in the MBD protein family

member, *MeCP2*, are the cause of Rett syndrome, a severe neurodevelopmental disorder (17) and MeCP2, MBD1, and MBD2 bind abnormal methylated DNA sites in human cancer cell lines (18-20).

Aberrant hypermethylation of CpG islands at the 5' end of tumor suppressor genes, leading to their transcriptional repression, seems to be a characteristic of human cancer cell lines and tumor tissues (18, 21, 22). A number of candidate genes approaches studies demonstrated conclusively that MBD proteins are directly involved in the epigenetic silencing of tumor suppressor genes by binding to aberrantly methylated regulatory regions (10, 19, 23-29). ChIP on chip experiments from limited number of hypermethylated silenced genes (7,776) indicate that among the 3 MBD proteins analyzed (*MeCP2*, MBD1, and MBD2), MBD2 exhibits the widest binding profile in human breast cancer cell lines (27). Moreover, in HeLa cells, a human uterine cervical cancer cell line, depletion of MBD proteins by RNA interference (RNAi) induces a release of transcriptional silencing of hypermethylated genes. In addition, the data indicate that for single MBD interference, MBD2 depletion is the protein most commonly involved in the observed release of gene silencing by far (30). All together, these studies are in favor of prominent role of MBD2 in silencing hypermethylated genes in cancer cell lines.

Data obtained from mouse models are also in favor of an important role of MBD2 in cancer cells. Strong reduction of intestinal adenomas has been observed in *Mbd2* knockout mice derived from *Apc*^{Min/-} mice (31), suggesting that loss of Mbd2 protects the mice from gene-hypermethylations associated with *Apc* mutation (32). In human cell lines, *MBD2* antisense inhibitors suppress tumorigenesis *in vitro* and *in vivo*, when these cells are implanted in nude mice as a model (33). The anti-tumoral action of MBD2 depletion is not due to intrinsic cellular toxicity, since morpholino, antisense RNA (33) or siRNA targeting MBD2 (~ 10 cell lines analyzed, unpublished data) affect neither cell survival nor growth cell

rate. Taken together, these data suggest that MBD2 protein is involved in the complex steps leading to the outgrowth of malignant cells in a new organ environment, and make MBD2 an attractive potential target for therapeutic intervention in cancer.

However, questions have been raised about the role of MBD proteins in gene silencing. Recently, ChIP on chip analysis of MeCP2 binding sites in neuroblastoma cell lines on tiling array (34) and identification of new partners (35) have suggested that MeCP2 may have a dual role and can activate silent genes. However, genome-wide mapping of MeCP2 binding sites in neuronal nuclei, using *MeCP2*^{-/-} mice as a control, indicates that, despite a very high abundance of MeCP2 molecules in neuronal cells, this protein selectively associates methylated regions and affect chromatin structure at global level (36). These data also suggest that the main function of MeCP2 is to reduced transcriptional noises in neuronal cells (36). Thus, further studies are needed in order to fully understand the role of MeCP2 in mental retardation. Some data also suggest that MBD2 may act as a gene activator. Yeast two-hybrid interaction screening has led to the discovery of two interacting proteins: TACC3 (transforming-acidcoiled-coil) (37) and MBD2in (38), and data obtained from methylated expression vectors indicate both proteins are able to relieve MBD2 repression (37). Moreover, association between MBD2 and a viral transcriptional stimulator (Tax) of the human T-cell leukemia virus type 1 (HTLV-1) has been reported (39). The authors have also found that the complex MBD2-Tax activates methylated expression vectors containing the HTLV-1 long terminal repeat as a promoter (39). A role of MBD2 in activation of unmethylated expression vectors has been also suggest from a study indicating that MBD2a (the long isoform of MBD2 transcripts) selectively interacts with RNA helicase A (RHA), a component of CREB transcriptional coactivator complexes (40).

Taken together, these data raise the question of a putative major role of MBD2 in transcriptional regulation in cancer cells. In order to address this point, we have analyzed the

characteristic of MBD2 binding sites in HeLa cells, a human cervix carcinoma cell line. Fragmented chromatin segments were selected by immunoprecipitation using antibodies directed against MBD2 and the corresponding DNA fragments were analyzed on high-density tiling array representing the transcription start site (TSS) regions (7.5 kb upstream and 2.5 kb downstream) of about 25,500 human promoters (ChIP on chip). The map of MBD2 binding sites were then compared with the map of methylated DNA sites and RNA polymerase II (POLR2) binding sites determined using methylated DNA precipitation (MeDP) and ChIP on chip approach, respectively. This genome wide analysis correlates MBD2 binding sites with DNA methylation and silent genes, indicating that MBD2 behaves like a true methylation-dependent transcriptional repressor in the human cancer HeLa cells.

MATERIALS AND METHODS

Cell culture

The cervical cancer cell line HeLa was obtained from ATCC (Rockville, MD) and grown in DMEM (Eagle, Sigma, L'isle d'Abeau, France) supplemented with 10% fetal bovine serum. Cells were grown at 37°C in a humidified 5% CO₂ atmosphere.

Recombinant 4xMBD proteins production and purification

pET-4xMBD and pET-4xMBD-R22A plasmids encoding in bacterial system recombinant wild type His₆-tagged 4xMBD and mutant His₆-tagged 4xMBD-R22A proteins, respectively, were kindly provided by Dr. A. Bird (41). Both recombinant proteins were purified, from induced *E. coli* Rosetta (DE3 pLysS) cultures, on Ni-NTA agarose (Qiagen, Courtaboeuf, France) using denaturation conditions and renaturation by stepwise dialysis.

Electrophoretic mobility shift assays

Electrophoretic mobility shift assays were performed using LightShift Chemiluminescent EMSA kit (Thermo Scientific, Rockford) in accordance to manufacturer's instructions. Briefly, 200 ng purified His₆-tagged 4xMBD protein in 1X binding buffer, 5 % glycerol, 5 mM MgCl₂, 50 ng/μL poly dI.dC, 0.05% NP-40, and 0.5 M Urea were pre-incubated 20 minutes in room temperature before the addition of 0.25 fmol biotin-5'end-labeled double-strand probes corresponding to the sequence ATC-AGA-CGT-TCG-CCG-GCG-GAT-TGG-CTT-GGC-TGC-GAA-GAA-GAT-A, either symmetrically methylated at CpG sites or

unmethylated (Eurogentec, Seraing, Belgium). After further 20 min incubation at room temperature, binding reactions were loaded on a 6% polyacrylamide gel in 0.5X TBE and ran at 100V for about 50 minutes. Binding reactions were then transferred (1 hour, 230 mA, 4°C) to nylon membrane and UV cross-linked for 12.5 minutes. Biotin-labeled probe was detected by chemiluminescence.

Methyl DNA precipitation (MeDP)

Fifteen μg of His₆-tagged 4xMBD proteins and 500 ng of HeLa DNA sheared to an average length of 300 - 600 bp were incubated 1 hour at room temperature in an interaction buffer containing 10 mM Tris pH 7.8, 50 mM KCl, 5 mM MgCl₂, 0.05% NP-40, 0.5 mM DTT, 0.5 M urea and 5% glycerol. Ni-NTA magnetic agarose beads (Qiagen) were added 1 hour. Then, beads were washed four times with the interaction buffer plus 0.3 M NaCl and eluted in 40 mM Tris pH 7.8, 10 mM EDTA, 0.5% SDS. A digestion with proteinase K (400 μg) was performed before purification of precipitated DNA by two phenol-chloroform extractions. This bound fraction was compared with input fraction, which corresponds to sonicated DNA that was only digested and purified. To test for enrichment of methylated sites, PCR amplifications of *BRCA1* (28), *hTERT* (29), *pS2* (42), *PARVG* (30), *RASSF1A* (30) and *GAPDH* promoters were performed on each MeDP samples. Six independent MeDP experiments were performed and DNA from input and bound fractions were collected.

Chromatin Immunoprecipitation (ChIP)

ChIP assays were carried out as previously described (28). Nucleoprotein complexes, sheared to an average length of 300 - 600 bp, were immunoprecipitated with 15 μl of two different

polyclonal anti-MBD2 antibodies (kindly provided by Dr. P. Wade and Dr. E. Ballestar); and 20 μ l of polyclonal anti-POLR2 (8WG16, Covance, California). The relative concentrations in the bound and input fractions of immunoprecipitated DNAs were assessed by real-time PCR amplification of *BRCA1* (28), *hTERT* (29), *pS2* (42), *PARVG* (30), *RASSF1A* (30), and *GAPDH* DNA segments. The sequences of the primers used are available upon request. Three independent ChIP experiments were performed for each antibody and DNA from input and bound fractions were collected and pooled. This procedure was performed two times to obtain 2 pools of input fraction and 2 pools of bound fraction for each antibody.

ChIP and MeDP on chip

The ChIP and MeDP DNAs from the input and bound fractions were amplified, labeled and hybridized on microarrays by ProfileXpert service (Lyon, France) according to Affymetrix™ protocols. Briefly, DNA fragments from ChIP and MeDP experiments were independently amplified by random PCR. The conservation of relative concentrations, in the bound and input fractions, DNA fragment precipitated by anti-MBD2, anti-POLR2 antibodies and 4XMBD was assayed by real-time PCR amplification of *BRCA1*, *hTERT*, *pS2*, *PARVG*, *RASSF1A*, and *GAPDH* promoters, on each ChIP or MeDP samples before and after amplification. The amplified DNAs, from each POLR2 ChIP and MeDP fractions, were pooled to obtain only one bound and one input fractions from, POLR2-ChIP or MeDP. MBD2 assays were performed in duplicates from 6 independent ChIP. For homogeneity with POLR2 and MeDP assays, fractions corresponded to 6 independent experiments were pooled to obtain 2 bound and 2 input fractions, for further analysis data obtained from MBD2-arrays were processed together, using the software described below. These 8 samples were then labeled using the GeneChip® WT Double - Stranded DNA Terminal Labelling Kit and hybridized to

the human tiling arrays (Human Promoter 1.0R Arrays), which were then washed and scanned.

Array analysis

Raw data from the scans were analyzed using Affymetrix® Tiling Analysis Software (TAS), using NCBI build 36/HG18 human reference genome, and the results were viewed in Affymetrix' Integrated Genome Browser (IGB) Software. ChIP and MeDP on chip

Reviewer's access to ArrayExpress :

Username: Reviewer_E-MEXP-2634

Password: 1269266717622

Data were analyzed using CisGenome software (www.biostat.jhsph.edu/~hji/cisgenome) and Ingenuity Pathway Analysis software (Redwood City, CA, USA). Statistical analyses were performed using Wolfram Mathematica software and Microsoft Excel software. Lists of methylated DNA regions, MBD2, and POL2R binding sites are available upon request.

RESULTS

Identification of MBD2 target gene promoters in HeLa Cells

To identify MBD2 target gene promoters, we selected DNA fragments (about 300 bp in length) associated with MBD2 by chromatin immunoprecipitation (ChIP) and subjected that to high-resolution ChIP on chip analysis.

ChIP experiments were processed from HeLa cross-linked chromatin as described in “Materials and Methods” section and previously (28, 29). Two different antibodies directed against MBD2 were used in three replicate experiments. Following initial testing of chromatin quality (data not shown), the specificity of the antibodies was confirmed by a set of genes (*BRCA1*, *pS2/TFF1*, and *hTERT*) previously analyzed in our own laboratory (28, 29) and data (*PARVG*, *RASSF1A*, and *GAPDH*) previously published (30) (**Fig. 1**). Next, we analyzed the immunoprecipitated DNAs on Affymetrix Human Promoter 1.0R Array comprising the kb -7.5 to +2.45 regions of around 25,500 human promoters. Immunoprecipitated DNAs and Input control DNAs were distributed into independent pools and separately amplified before hybridization on arrays. Control experiments indicated that the PCR amplification step had retained the specificity of the starting ChIP samples (data not shown).

To determine peak detection parameters, we used, as a training set, the data previously obtained during control experiments of PCR amplification step for *BRCA1*, *pS2*, *RASSF1A*, *GAPDH*, *hTERT*, and *PARVG*. We visualized the array signals using the integrated genome browser (IGB) and we fitted the thresholds at $\log_2 > 0.3$. Finally, we validated new regions identified with this threshold value using data published for MBD2 sites from an independent laboratory (19). We found a good concordance (78%, 14/18) between MBD2 sites independently identified by direct ChIP assays in HeLa cells (19) and our own array analyses

(**supplementary data, Table S1**). This result demonstrates the validity of our ChIP on chip approach, the threshold value, $\log_2 > 0.3$, was used for further the analysis of array data.

Using this strategy, we identified 8,257 gene promoters bound by MBD2 among the 23,054 autosome promoter regions represented on the array. In this analysis the probes corresponding to the promoter regions of the chromosome X were excluded, since HeLa cells were derived from a female donor and contain three copies of chromosome X (43), and, this abnormal copy number may induce bias, due to X inactivation, in the determination of its methylation patterns.

Characterization of the MBD2 binding regions identified by ChIP on chip based on their relative positions to TSSs

In order to gain insight on the localization of the signals corresponding to MBD2 binding sites, we analyzed their distributions relative to the transcription starting sites. This study was performed using the chromosome 1 as a model. This chromosome was chosen since it contains the greatest number of genes. The distribution of the signals were determined by the ratio between the number of “positive” probes *versus* the number of the probes present on the array in the corresponding regions, since the probes were not equally distributed on the arrays. Therefore these percentages were relatively independent of the threshold used. The values were then normalized using the median value for each interval and plotted against the distance to the TSS at a 50 bp resolution (**Fig. 2**).

This genome-wide averaging analysis (raw data are shown in supplementary data **Fig. S1**) revealed enrichment of MBD2 binding sites that peaked in the TSS (Fig. 2). These data indicated that MBD2 binding sites are preferentially localized at TSS regions. The positive signal started to increase at a distance ~ 700 bp upstream of the TSS, and exhibited maximum

values between 0 and ~200 bp downstream of the TSS. This region (from position -700 to +200) was, thus, used for further analysis.

Analysis of promoters DNA Methylation profile

Raising the question of a putative major role of MBD2 in DNA methylation-dependant transcriptional repression, we first examined the distribution of DNA methylation among the promoters. To this aim, we performed a Methylation DNA precipitation (MeDP) on chip using a recombinant 4xMBD protein to selectively precipitate the methylated sequences.

The recombinant 4xMBD protein contains four Methyl-CpG-Binding Domain (MBD) cloned from the Methyl-CpG-Binding Domain protein 1 (MBD1). This protein associates methylated double-stranded (ds) DNA and binds ds-oligonucleotides with high affinity (0.5×10^{-6} M to 0.2×10^{-7} M, depending on the number of methylated CpG) (41). Thus, the 4xMBD represents an interesting tool for selecting methylated DNA fragments.

The specificity of this protein for methylated CpGs, in our own experimental conditions, was determined by band-shift assays using a double-stranded 43-bp probe, containing 5 CpGs. The methylated oligonucleotide was efficiently recognized by the 4xMBD protein, while the 4xMBD protein did not interact with the unmethylated probe (**Fig. 3**). Moreover, negative controls were performed with the 4xMBD-R22A protein which carries a R22A point mutation that disrupts its DNA binding activity (41). As expected, the mutated recombinant protein interacted neither with the methylated nor with the unmethylated probes (Fig. 3).

As this recombinant protein was for the first time used to identify methylated regions at the genomic scale, we further tested its functionality to MeDP. Sonicated genomic HeLa DNA was incubated with 4xMBD proteins and DNA fragments associated with these proteins

were purified. The relative amounts of a set of genes, previously analyzed by bisulfite sequencing (19, 28, 29), were determined by quantitative PCR (Q-PCR). Next, we investigated the potential effect of CpG density on the interaction between the 4xMBD and methylated genomic DNA fragments. We chose three methylated DNA fragments: a segment of the *BRCA1* 5' end embedded in a CpG island (CGi) ($\text{CpG}_{\text{obs/exp}} = 0.65$; G+C = 57%), a segment of a non-CGi *PARVG*, and a segment of *pS2* 5' end not included in a CGi. Three unmethylated DNA segments were also quantified: a segment including the exon 1 of *BRCA1* (also included in the *BRCA1* CGi), a segment of the 5' end of *GAPDH* included in a CGi ($\text{CpG}_{\text{obs/exp}} = 0.72$; G+C = 67.4%), and a segment of the 5' end of *RASSF1A* CGi ($\text{CpG}_{\text{obs/exp}} = 0.79$; G+C = 62%). Q-PCR analysis (Fig. 1) indicated a significant enrichment in methylated DNA segments (*BRCA1*, $p = 0.0017$; *PARVG*, $p = 0.0028$; and *pS2/TFF1*, $p = 0.02$) and a significant depletion in unmethylated DNA sequences (*BRCA1* exon1, $p = 0.002$; *GAPDH*, $p = 0.02$; *RASSF1A*, $p = 0.01$), in the bound fractions. Moreover, CpG density does not seem to affect the selection efficiency of methylated DNA fragments. These data support that MeDP technique is highly specific and sensitive, thus selected methylated DNA fragments were hybridized to Affymetrix Promoter Arrays to characterize DNA methylation profiles of promoters at the genomic scale.

A similar approach to MBD2 CHIP on chip was used to analyze data from MeDP experiments. The threshold, $\log_2 > 0.1$, to provide the expected negative and positive regions, was also fitted with the six genes, *BRCA1*, *pS2*, *RASSF1A*, *GAPDH*, *hTERT*, and *PARVG*, previously used as control. For validation, data obtained from methylation sensitive restriction enzyme (*CfoI*) digestion was compared to MeDP on chip data. Among the 27 genes exhibiting a methylated 5' end, determined from the method based on a Multiplex Ligation-dependent Probe Amplification protocol (MLPA) (**supplementary data, Table S2**), 77% of them were also found methylated using the MeDP on chip approach.

To accurately describe the distribution of DNA methylation among the 23,054 gene promoters represented on the Affymetrix array, the percentage of probes hybridizing to DNA fragments selectively precipitated with the recombinant 4xMBD protein was normalized by CpG density at a 200 bp resolution. DNAs were randomly fragmented by sonication to a median length of 300 bp. Thus, if the majority of CpGs were methylated in fragments selected by the 4xMBD protein in an interval containing several probes, the percentage of probes hybridizing with the bound fractions should be proportional to the CpG density. To obtain a reliable estimation of the methylated CpG density, the ratio CpGs observed *versus* CpGs expected was determined for each set of probes. We choose a 200bp resolution interval because it corresponds to about 6 probes, which can therefore hybridize to different DNA fragments. We observed that the relative DNA methylation gradually decreased when approaching the TSS regions both upstream and downstream and reached its minimum value at the TSS position (**Fig. 4**). Therefore, this analysis suggests that the TSS regions of chromosome 1 is hypomethylated as it was observed at the TSS regions of chromosomes 6, 20, 21, and 22 of normal and transformed cells (44, 45).

Moreover, analysis of single-base-resolution maps of methylated cytosines (46) indicated that the number of genes exhibiting a methylated DNA sequence is similar in HeLa and in IMR90 cells. For example, MeDP on chip experiments from HeLa cells identified 969 promoter regions on chromosome 19 that possess a methylated sequence, while, in IMR90, high throughput sequencing indicated that 814 TSS regions (from position -700 to +200) exhibited sequences (mean length of 286 bp \pm 130) containing at least 4.01 methylated CpGs per 100 bp. Since MeDP on chip and high throughput sequencing are not based on the same methods, this comparison indicates that MeDP on chip is a reliable method for the identification of potential methyl-DNA binding protein sites. Thus, in this analysis a “methylated gene” is a gene exhibiting, at least, a 200 bp DNA sequence in length containing

about 8 (or more) methylated CpGs located in the -700 to +200 region, relative to their TSS.

Correlation of MBD2 binding sites with DNA Methylation

As expected for a methyl-CpG-binding protein, the majority of the MBD2 binding sites (74%; 6,108 out of 8,257) associated genes methylated at their 5' end (from position -700 to +200) (**Fig. 5A**). Statistic analysis indicated that this association was not the result of a random process (χ^2 test = 1.1×10^{-145}) suggesting a specific recruitment of MBD2 at methylated sites.

The majority of studies concentrated their efforts to describe the MBD proteins occupancy on hypermethylated CpG island as these promoters seem to be attractive substrates in cancer cells on which to exert their function. We identified genes exhibiting a CpG island at TSS regions (from position -700 to +200) using the criteria defined by Takai and Jones (regions of DNA of greater than 500 bp with a G+C equal to or greater than 55% and observed CpG versus expected CpG of 0.65 or greater (47)), the data from USCS data bank and direct analysis of bound-fragments using CpGProD software (<http://pbil.univ-lyon1.fr/software/cpgprod.html>). The analysis of ChIP and MeDP on chip experiments performed in HeLa cells indicated that 13.6% (3143 out of 23055) of the promoters exhibited both a methylated-MBD2 positive region and a CpG island at their 5' end. However, these characteristics were also observed for about the same percentage (12.8%) of genes without CpG island, suggesting that the presence of a CpG island is not a discriminating factor for the binding of this epigenetic factor, at least in HeLa cells.

Motif analysis reveals no specific MBD2 binding sequence motifs excepted the presence of at least one CpG site

ChIP analysis has suggested that MBD2 proteins may possess some specificity in their association with methylated promoters (27, 28). However, this issue remains an open question in the field. Moreover, in contrast to MeCP2 that needs A/T sequences adjacent to methyl-CpG (48), sequence analysis of 435 cloned MBD2-ChIP fragments have suggested that MBD2 binds methylated CpG independently of the surrounding sequences (48). To obtain further insights of the potential sequence specificity of MBD2 binding, we searched for common motifs from binding fragments identified by MBD2-ChIP on chip experiments. Using Scope motif finder web site (<http://genie.dartmouth.edu/scope/>), we analyzed the 7,763 MBD2 binding sequences (length-median-value of 184 bp) identified in the promoter regions and determined the top 20 common motifs exhibiting at least 6 bp (**Table 1**). Data obtained indicated that the presence of a CpG is a constant motif among these consensus sequences, while no other constant motifs were observed, with, maybe, the exception of a CCG sequence that seems overrepresented among the common motifs identified (Table 1). Taken together, these findings indicate that MBD2 acts primarily as methyl-CpG-binding protein and suggest that the flanking sequences have little effect on MBD2 binding.

Functional consequences of DNA methylation and MBD2 binding on gene regulation

MBD2 is known to mediate DNA methylation gene silencing. However, so far this observation was mainly based on experiments with individual or selected set of tumor suppressor genes. Moreover, some recently data also suggest that MBD2 may act as a gene activator (37, 38, 40, 49). To directly assess the functional consequences of MBD2 binding on gene transcription with a genome-wide basis, we performed ChIP on chip using an antibody against RNA polymerase II (POLR2) in parallel to MBD2-ChIP on chip and MeDP on chip. In order to identify silent genes we used the absence of POLR2 binding as an indicator of the

transcriptional activity absence (Fig. 1). The list of silent genes identified by the $\log_2 < -0.1$ threshold was compared with the analysis of HeLa cell transcripts using CodeLink microarray (**supplementary data, Table S3**). Among this list, 5,813 genes were unambiguously identified on the list generated by the transcriptomic method. We found that the transcripts corresponding to 4,695 of them (80% of concordance) were not detected using this latter method. We have then established a reliable approach to identify a potential relationship between MBD2 binding sites, DNA methylation and transcriptionally inactive promoters.

Among the genes methylated at their 5' end (from position -700 to +200), 71% (10,370 out of 14,532) also exhibited a POLR2 free 5' end (Fig. 5B). This significant association (χ^2 test = 7.7×10^{-212}) was in line with many studies reporting an inverse correlation between DNA methylation and gene expression. Furthermore, more than 70% of POLR2 negative regions were methylated, suggesting that DNA methylation is a crucial event in gene silencing.

The list of MBD2-positive genes was also compared with genes exhibiting a POLR2 free TSS region (Fig. 5C). A non-random association between MBD2-positive genes and POLR2 free genes was observed ($\chi^2 = 7.7 \times 10^{-3}$). Indeed, the majority of MBD2-positive genes (65%, 5,365/8,273) were silenced, as determined by the absence of POLR2, suggesting that MBD2 is involved in the transcriptional repression of these genes.

Finally, when data obtained from the analysis of the regions (from position -700 to +200) surrounding the TSSs were compared, we have found that 51% of MBD2-positive genes were also methylated and silenced. These data indicate a non-random association (multiple correspondence analysis; $p < 0.0001$) between these parameters (Fig. 5D) and from the 23,054 promoter regions represented on the microarray, about 18% of them exhibit MBD2-positive, methylated, and POLR2 free TSS regions. Among the 23,054 putative promoters analyzed in this study, 12,486 have been annotated with high-confidence (50),

using three different data banks. When the analysis was restricted to this subset of promoters, data obtained (supplementary data, Fig. S2) did not significantly modify the significance of these associations. All together, these data supports the dominant model of MBD2 function as that of methylation-dependant transcriptional silencer.

Relationship between the distance of MBD2 binding sites from transcription start sites and gene silencing

Although both MeCP2 and MBD1 exhibit a long-range effect on gene expression, the efficiency of the transcriptional repression mediated by the transcriptional repression domains (TRD) of MeCP2 and MBD1 is differentially affected by the distance from TSS of reporter genes carried by expression vectors (3, 51). The strength of the repression driven by the MBD1-TRD is not affected by the distance even when it is lays more than 2,000 bp from the TSS (51), while the repressive effect of MeCP2-TRD progressively declines as the distance from TSS in the regions analyzed increases (-4,000 to -2,200 bp) (52).

In the course of these experiments, we have mapped the MBD2 binding sites form large regions (10 kb) around TSSs and we noticed that some genes (1,008) exhibited only one MBD2 positive region at 900 pb resolution. We therefore investigated the effect of the distance from TSSs on repression linked to MBD2. The number of genes deprived of POLR2 at their TSS and exhibiting a methylated region bound by MBD2, was plotted against the distance of MBD2 binding sites from their TSS (**Fig. 6**). The numbers of silent genes lessened with distance, we identified 63 and 186 silent genes associated with MBD2 binding located in the intervals, -1,600 to -2,500 and +1,100 to +2,000, respectively, while 350 were deprived in POLR2 when MBD2 binding sites were located in the -700 to +200 region. Interestingly, a distance-effect was also observed from MBD2 binding sites lying downstream the TSS, 246

and 163 silenced genes exhibited MBD2 binding sites in the +200 to +1,100 and +1,100 to +2,000 regions, respectively. These data indicate that MBD2 binding sites near the TSSs have a predominant role in gene silencing.

MBD2 preferentially associate genes involved in cell death and cell cycle in HeLa cells

The biological functions of silent genes exhibiting methylated 5' end region and bound by MBD2 (-700 to +200 interval) were analyzed using Ingenuity Pathway Analysis software. Categorization by molecular and cellular function of these genes leads to 3 majors groups (**Table 2**), (i) gene expression, (ii) cell death, and (iii) cell cycle. Among genes categorized in these 3 functional groups, lot of tumor-associated genes well known to be hypermethylated in cancer cells, such as *MGMT*, *p15^{INK4B}*, *DAPK*, Cyclin D2 and *MMP*, were found (53, 54). Therefore, among the 62 categories analyzed, the emerging groups are directly related to biological functions altered in cancer cells, suggesting that MBD2 is involved in the neoplastic process.

DISCUSSION

Evidences for involvement of proteins possessing a methyl-CpG-binding domain in gene silencing have been obtained from a large variety of cell lines (18). Among the MBD proteins MBD2 seems to be an important actor in silencing of aberrantly methylated genes in cancer cells. However, recent data suggest a dual role for these MBD proteins (37). Several experiments, using different approaches (identification of interacting proteins, expression of transfected vectors...) have suggested that MBD2 may act as a gene activator (39, 49). In order to define the major role of MBD2, at whole genome level, we have analyzed the characteristics of MBD2 binding sites, in the HeLa cancer cell line.

Distribution of MBD2 binding sites on the promoter regions

MBD2, DNA methylation, and POLR2 sites were mapped on array representing more than 25,500 promoter regions. Using chromosome 1 as a model, the position of MBD2 binding sites relative to TSS were determined. The percentage of probes hybridized to DNA fragments selected by ChIP assays, using antibodies directed against MBD2, reached a maximum value at TSSs, suggesting that MBD2 is involved in the regulation of the transcription initiation process.

MBD2 binding sites were also observed throughout the regions (position from -7,500 to +2,450) represented on the array. Recent study of *GATA-4* locus has shown that MBD2 may participate to long-range chromatin interaction mediated by loops structure and lead to gene silencing (55). Thus, we cannot exclude that some of these MBD2 binding sites are involved in this transcriptional regulation process.

Nevertheless, MBD2 binding sites peak at the TSS regions, thus, in order to investigate the relationship between MBD2 binding, DNA methylated regions, and absence of POLR2, we focused our analysis on the -700 to +200 interval, corresponding to MBD2 peak.

Non-random association between MBD2 binding and DNA methylation

The large majority (74%) of genes bound by MBD2 exhibited methylated sites. Thus, at whole genome level the correlation between DNA methylation and MBD2 binding sites was observed. We also found MBD2 binding sites in DNA methylation free regions, and we cannot exclude that some unmethylated DNA sequences were targeted by MBD2. Nevertheless, it should be noted that array data were analyzed using TAS software that excluded potentially non-significant signals. For example, 1 or 2 positive probes are not taken in account if they are surrounded by negative probes. Although this analysis leads to a reliable identification of positive regions, some DNA fragments with only one methylated CpG may be missed, whereas MBD2 can associate DNA sequences containing only one methylated CpG (4). Taken together, these data indicate that the major mechanism involved in the association of MBD2 with DNA is the methylation of the targeted regions. Search for consensus sequences from the MBD2- binding sites failed to identify common motifs among these sites, with the exception of the presence of at least one CpG site. This observation is also in favor of a MBD2 specificity for methylated sites.

MBD2 associates promoter of silent genes in DNA methylation-dependent manner

The absence of POLR2 does not mark all silent genes, as POLR2 needs to be assembled in a complex prior initiating transcription (56). Nevertheless, comparison between TAF-1 (TATA

box binding protein associated factor) and POLR2-bound promoters have shown a marked concordance between POLR2 and active transcription (50). We choose this criterion, absence of POLR2, for two reasons. First of all, the comparison between MBD2 presence and the absence of POLR2 avoids potential biases due to hybridization kinetics and/or probes present on the array, since we compared a positive signal with a negative signal. Secondly, the absence of POLR2 unambiguously indicates that the gene is not transcribed at the sites analyzed. This analysis of POLR2 binding sites confirmed the correlation between DNA methylation and gene silencing, 72% of methylated genes were deprived in POLR2 and 72% of silent genes exhibited methylated sequences in the -700 to +200 region, suggesting that DNA methylation is an important process in transcriptional repression.

As expected from the data described above, MBD2 binding sites were non-randomly distributed ($p < 0.0001$) among the methylated and the silent genes. The percentage (51%) of genes exhibiting MBD2 binding sites, methylated at their 5' ends and silent, may be under-evaluated, since the criterion used for the identification of untranscribed genes is the absence of POLR2.

Therefore, genome-wide analysis of MBD2 binding sites indicates that this methylation-dependent DNA binding protein acts as a transcriptional repressor of genes exhibiting methylated DNA regions at their promoter regions.

Distance of MBD2 binding sites, from TSSs, lessened the strength of association between the presence MBD2 and gene silencing

We observed a significant association between MBD2 binding, DNA methylation and gene silencing, in TSS regions (from position -700 to +200). The number of silent genes associated with MBD2-methylated regions gradually decreased with distance of positive MBD2-

methyated regions from TSSs. These data suggest that the strength of transcriptional repression mediated by MBD2 is lessened with distance. The effect of distance on the repressive activity of MBD proteins seems to be MBD-specific, MBD1 can repress vectors expression even when relatively far from the TSS, while MeCP2 needs to be close to the upstream regions of reporter genes (3, 51). It should be noted that the distance-effect on MBD protein activity is not the result of direct comparison, since data for MBD1 and MeCP2 have been obtained from transient transfection and the consequence of MBD1 and MeCP2 binding at the downstream regions of TSSs have not been investigated. The association between MBD2 binding and silent gene was predominantly observed at the TSS regions, and, therefore, suggests that MBD2 binding at these regions inhibits transcriptional initiation. Nevertheless, we cannot exclude that MBD2 binding at downstream region from the TSS alters transcriptional initiation and/or elongation. In line with this hypothesis, it was shown that methylation-free regions downstream the TSSs is required for efficient initiation (57) and intragenic methylation strongly reduces the efficiency of POLR2 elongation (58).

MBD2, a major actor in methylation-dependent gene silencing

Among the silent genes, about 42% of them exhibited methylated sites in the -700 to +200 region and were also bound by MBD2, suggesting that MBD2 is predominant in the transcriptional repression mediated by DNA methylation, in HeLa cells. Moreover, 41% of the methylated and silent promoters were also targeted by MBD2 indicating that this protein accounts for almost half of the genes silenced in a methylation-dependent manner. This value is relatively high, since the other MBD proteins, MBD1 and MeCP2, were also expressed in HeLa cells (29) and methylation-dependent repressors, such as Kaiso, were also present (5).

The analysis of gene transcription also indicated that MBD2 play a major role in gene

silencing. In mouse-tail fibroblast cell lines lacking *Mbd2*, *in vitro* methylated expression vectors are inefficiently repressed, while *Mecp2* deletion does not affect the repression of transiently transfected methylated templates (59). Using an array representing a limited number of genes (7,327 cDNA sequences), a release of transcriptional silencing upon depletion of MBD proteins by RNAi was observed and MBD2 depletion was the protein most commonly involved in gene re-activation in HeLa cells (30). Furthermore, the list of genes targeted by MBD2, revealed by ChIP on chip experiments, covers most of the disrupted molecular cellular functions in cancer cells, indicating that MBD2 is an important parameter in the maintenance of their transcriptional silencing.

Thus, genome-wide analysis and experimental manipulation of MBD2 levels are in favor of a predominant role of MBD2 in methylation-dependent gene silencing in epithelial cancer cells.

ACKNOWLEDGMENTS

We are most grateful to Dr Adrian Bird for the pET-4xMBD and pET-4xMBD-R22A vectors.

The present work was supported by the Ligue Nationale contre le Cancer (Comité du Rhône) and the Institut National contre le Cancer.

REFERENCES

1. Bird, A., and Tweedie, S. (1995) Transcriptional noise and the evolution of gene number. *Philos Trans R Soc Lond B Biol Sci* **349**, 249-253
2. Bird, A. P., and Wolffe, A. P. (1999) Methylation-induced repression--belts, braces, and chromatin. *Cell* **99**, 451-454
3. Nan, X., Campoy, F. J., and Bird, A. (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88**, 471-481
4. Hendrich, B., and Bird, A. (1998) Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol Cell Biol* **18**, 6538-6547
5. Sasai, N., and Defossez, P. A. (2009) Many paths to one goal? The proteins that recognize methylated DNA in eukaryotes. *Int J Dev Biol* **53**, 323-334
6. Sakai, H., Urano, T., Ookata, K., Kim, M. H., Hirai, Y., Saito, M., Nojima, Y., and Ishikawa, F. (2002) MBD3 and HDAC1, two components of the NuRD complex, are localized at Aurora-A-positive centrosomes in M phase. *J Biol Chem* **277**, 48714-48723
7. Dhasarathy, A., and Wade, P. A. (2008) The MBD protein family-reading an epigenetic mark? *Mutat Res* **647**, 39-43
8. Wade, P. A., Gogonne, A., Jones, P. L., Ballestar, E., Aubry, F., and Wolffe, A. P. (1999) Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nat Genet* **23**, 62-66
9. Majumder, S., Kutay, H., Datta, J., Summers, D., Jacob, S. T., and Ghoshal, K. (2006) Epigenetic regulation of metallothionein-i gene expression: differential regulation of

- methylated and unmethylated promoters by DNA methyltransferases and methyl CpG binding proteins. *J Cell Biochem* **97**, 1300-1316
10. Kondo, E., Gu, Z., Horii, A., and Fukushige, S. (2005) The thymine DNA glycosylase MBD4 represses transcription and is associated with methylated p16(INK4a) and hMLH1 genes. *Mol Cell Biol* **25**, 4388-4396
 11. Hendrich, B., Hardeland, U., Ng, H. H., Jiricny, J., and Bird, A. (1999) The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature* **401**, 301-304
 12. Barr, H., Hermann, A., Berger, J., Tsai, H. H., Adie, K., Prokhortchouk, A., Hendrich, B., and Bird, A. (2007) Mbd2 contributes to DNA methylation-directed repression of the Xist gene. *Mol Cell Biol* **27**, 3750-3757
 13. Fournier, C., Goto, Y., Ballestar, E., Delaval, K., Hever, A. M., Esteller, M., and Feil, R. (2002) Allele-specific histone lysine methylation marks regulatory regions at imprinted mouse genes. *Embo J* **21**, 6560-6570
 14. Zhao, X., Ueba, T., Christie, B. R., Barkho, B., McConnell, M. J., Nakashima, K., Lein, E. S., Eadie, B. D., Willhoite, A. R., Muotri, A. R., Summers, R. G., Chun, J., Lee, K. F., and Gage, F. H. (2003) Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function. *Proc Natl Acad Sci U S A* **100**, 6777-6782
 15. Chen, R. Z., Akbarian, S., Tudor, M., and Jaenisch, R. (2001) Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nat Genet* **27**, 327-331
 16. Guy, J., Hendrich, B., Holmes, M., Martin, J. E., and Bird, A. (2001) A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat Genet* **27**, 322-326

17. Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., and Zoghbi, H. Y. (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* **23**, 185-188
18. Esteller, M. (2007) Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet* **8**, 286-298
19. Lopez-Serra, L., Ballestar, E., Fraga, M. F., Alaminos, M., Setien, F., and Esteller, M. (2006) A profile of methyl-CpG binding domain protein occupancy of hypermethylated promoter CpG islands of tumor suppressor genes in human cancer. *Cancer Res* **66**, 8342-8346
20. Lin, X., and Nelson, W. G. (2003) Methyl-CpG-binding domain protein-2 mediates transcriptional repression associated with hypermethylated GSTP1 CpG islands in MCF-7 breast cancer cells. *Cancer Res* **63**, 498-504
21. Baylin, S. B., and Ohm, J. E. (2006) Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction? *Nat Rev Cancer* **6**, 107-116
22. Ehrlich, M., Jiang, G., Fiala, E., Dome, J. S., Yu, M. C., Long, T. I., Youn, B., Sohn, O. S., Widschwendter, M., Tomlinson, G. E., Chintagumpala, M., Champagne, M., Parham, D., Liang, G., Malik, K., and Laird, P. W. (2002) Hypomethylation and hypermethylation of DNA in Wilms tumors. *Oncogene* **21**, 6694-6702
23. Magdinier, F., and Wolffe, A. P. (2001) Selective association of the methyl-CpG binding protein MBD2 with the silent p14/p16 locus in human neoplasia. *Proc Natl Acad Sci U S A* **98**, 4990-4995
24. Nguyen, C. T., Gonzales, F. A., and Jones, P. A. (2001) Altered chromatin structure associated with methylation-induced gene silencing in cancer cells: correlation of accessibility, methylation, MeCP2 binding and acetylation. *Nucleic Acids Res* **29**, 4598-4606

25. Lin, X., Asgari, K., Putzi, M. J., Gage, W. R., Yu, X., Cornblatt, B. S., Kumar, A., Piantadosi, S., DeWeese, T. L., De Marzo, A. M., and Nelson, W. G. (2001) Reversal of GSTP1 CpG island hypermethylation and reactivation of pi-class glutathione S-transferase (GSTP1) expression in human prostate cancer cells by treatment with procainamide. *Cancer Res* **61**, 8611-8616
26. Koizume, S., Tachibana, K., Sekiya, T., Hirohashi, S., and Shiraishi, M. (2002) Heterogeneity in the modification and involvement of chromatin components of the CpG island of the silenced human CDH1 gene in cancer cells. *Nucleic Acids Res* **30**, 4770-4780
27. Ballestar, E., Paz, M. F., Valle, L., Wei, S., Fraga, M. F., Espada, J., Cigudosa, J. C., Huang, T. H., and Esteller, M. (2003) Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer. *Embo J* **22**, 6335-6345
28. Auriol, E., Billard, L. M., Magdinier, F., and Dante, R. (2005) Specific binding of the methyl binding domain protein 2 at the BRCA1-NBR2 locus. *Nucleic Acids Res* **33**, 4243-4254
29. Chatagnon, A., Bougel, S., Perriaud, L., Lachuer, J., Benhattar, J., and Dante, R. (2009) Specific association between the methyl-CpG-binding domain protein 2 and the hypermethylated region of the human telomerase reverse transcriptase promoter in cancer cells. *Carcinogenesis* **30**, 28-34
30. Lopez-Serra, L., Ballestar, E., Ropero, S., Setien, F., Billard, L. M., Fraga, M. F., Lopez-Nieva, P., Alaminos, M., Guerrero, D., Dante, R., and Esteller, M. (2008) Unmasking of epigenetically silenced candidate tumor suppressor genes by removal of methyl-CpG-binding domain proteins. *Oncogene* **27**, 3556-3566

31. Sansom, O. J., Berger, J., Bishop, S. M., Hendrich, B., Bird, A., and Clarke, A. R. (2003) Deficiency of Mbd2 suppresses intestinal tumorigenesis. *Nat Genet* **34**, 145-147
32. Eads, C. A., Nickel, A. E., and Laird, P. W. (2002) Complete genetic suppression of polyp formation and reduction of CpG-island hypermethylation in *Apc*(Min/+) *Dnmt1*-hypomorphic Mice. *Cancer Res* **62**, 1296-1299
33. Campbell, P. M., Bovenzi, V., and Szyf, M. (2004) Methylated DNA-binding protein 2 antisense inhibitors suppress tumourigenesis of human cancer cell lines in vitro and in vivo. *Carcinogenesis* **25**, 499-507
34. Yasui, D. H., Peddada, S., Bieda, M. C., Vallero, R. O., Hogart, A., Nagarajan, R. P., Thatcher, K. N., Farnham, P. J., and Lasalle, J. M. (2007) Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. *Proc Natl Acad Sci U S A* **104**, 19416-19421
35. Chahrour, M., Jung, S. Y., Shaw, C., Zhou, X., Wong, S. T., Qin, J., and Zoghbi, H. Y. (2008) MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* **320**, 1224-1229
36. Skene, P. J., Illingworth, R. S., Webb, S., Kerr, A. R., James, K. D., Turner, D. J., Andrews, R., and Bird, A. P. (2010) Neuronal MeCP2 Is Expressed at Near Histone-Octamer Levels and Globally Alters the Chromatin State. *Mol Cell* **37**, 457-468
37. Angrisano, T., Lembo, F., Pero, R., Natale, F., Fusco, A., Avvedimento, V. E., Bruni, C. B., and Chiariotti, L. (2006) TACC3 mediates the association of MBD2 with histone acetyltransferases and relieves transcriptional repression of methylated promoters. *Nucleic Acids Res* **34**, 364-372
38. Lembo, F., Pero, R., Angrisano, T., Vitiello, C., Iuliano, R., Bruni, C. B., and Chiariotti, L. (2003) MBDin, a novel MBD2-interacting protein, relieves MBD2

- repression potential and reactivates transcription from methylated promoters. *Mol Cell Biol* **23**, 1656-1665
39. Ego, T., Tanaka, Y., and Shimotohno, K. (2005) Interaction of HTLV-1 Tax and methyl-CpG-binding domain 2 positively regulates the gene expression from the hypermethylated LTR. *Oncogene* **24**, 1914-1923
 40. Fujita, N., Watanabe, S., Ichimura, T., Tsuruzoe, S., Shinkai, Y., Tachibana, M., Chiba, T., and Nakao, M. (2003) Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression. *J Biol Chem* **278**, 24132-24138
 41. Jorgensen, H. F., Adie, K., Chaubert, P., and Bird, A. P. (2006) Engineering a high-affinity methyl-CpG-binding protein. *Nucleic Acids Res* **34**, e96
 42. Chatagnon, A., Ballestar, E., Esteller, M., and Dante, R. (2010) A Role for Methyl-CpG Binding Domain Protein 2 in the Modulation of the Estrogen Response of pS2/TFF1 Gene. *PLoS One* **5**, e9665
 43. Macville, M., Schrock, E., Padilla-Nash, H., Keck, C., Ghadimi, B. M., Zimonjic, D., Popescu, N., and Ried, T. (1999) Comprehensive and definitive molecular cytogenetic characterization of HeLa cells by spectral karyotyping. *Cancer Res* **59**, 141-150
 44. Zhang, Y., Rohde, C., Tierling, S., Jurkowski, T. P., 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 Genet* **5**, e1000438
 45. Eckhardt, F., Lewin, J., Cortese, R., Rakyan, V. K., Attwood, J., Burger, M., Burton, J., Cox, T. V., Davies, R., Down, T. A., Haefliger, C., Horton, R., Howe, K., Jackson, D. K., 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. *Nat Genet* **38**, 1378-1385
46. Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., Nery, J. R., Lee, L., Ye, Z., Ngo, Q. M., Edsall, L., Antosiewicz-Bourget, J., Stewart, R., Ruotti, V., Millar, A. H., Thomson, J. A., Ren, B., and Ecker, J. R. (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315-322
47. Takai, D., and Jones, P. A. (2002) Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci U S A* **99**, 3740-3745
48. Klose, R. J., Sarraf, S. A., Schmiedeberg, L., McDermott, S. M., Stancheva, I., and Bird, A. P. (2005) DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. *Mol Cell* **19**, 667-678
49. Fujita, H., Fujii, R., Aratani, S., Amano, T., Fukamizu, A., and Nakajima, T. (2003) Antithetic effects of MBD2a on gene regulation. *Mol Cell Biol* **23**, 2645-2657
50. Weber, M., Hellmann, I., Stadler, M. B., Ramos, L., Paabo, S., Rebhan, M., and Schubeler, D. (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* **39**, 457-466
51. Ng, H. H., Jeppesen, P., and Bird, A. (2000) Active repression of methylated genes by the chromosomal protein MBD1. *Mol Cell Biol* **20**, 1394-1406
52. Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386-389
53. Mulero-Navarro, S., and Esteller, M. (2008) Epigenetic biomarkers for human cancer: the time is now. *Crit Rev Oncol Hematol* **68**, 1-11

54. Muller, I., Wischniewski, F., Pantel, K., and Schwarzenbach, H. Promoter- and cell-specific epigenetic regulation of CD44, Cyclin D2, GLIPR1 and PTEN by Methyl-CpG binding proteins and histone modifications. *BMC Cancer* **10**, 297
55. Tiwari, V. K., McGarvey, K. M., Licchesi, J. D., Ohm, J. E., Herman, J. G., Schubeler, D., and Baylin, S. B. (2008) PcG proteins, DNA methylation, and gene repression by chromatin looping. *PLoS Biol* **6**, 2911-2927
56. Sandelin, A., Carninci, P., Lenhard, B., Ponjavic, J., Hayashizaki, Y., and Hume, D. A. (2007) Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nat Rev Genet* **8**, 424-436
57. Appanah, R., Dickerson, D. R., Goyal, P., Groudine, M., and Lorincz, M. C. (2007) An unmethylated 3' promoter-proximal region is required for efficient transcription initiation. *PLoS Genet* **3**, e27
58. Lorincz, M. C., Dickerson, D. R., Schmitt, M., and Groudine, M. (2004) Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nat Struct Mol Biol* **11**, 1068-1075
59. Hendrich, B., Guy, J., Ramsahoye, B., Wilson, V. A., and Bird, A. (2001) Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. *Genes Dev* **15**, 710-723

FIGURE LEGENDS

Figure 1. Efficiency and specificity of ChIP and MeDP assays. After sonication, DNAs or cross-linked chromatin were, respectively, precipitated with 4xMBD recombinant proteins in MeDP assays or with anti-MBD2 or anti-POLR2 antibodies in ChIP assays. To determine the efficiency and specificity of the assays in order to use them for microarray analysis, 7 regions were analyzed by quantitative PCR. These 7 regions were chosen to represent a panel of combination between methylated/unmethylated and MBD2- and POLR2- bound/unbound CGi/non-GGi regions. For each experiment and each region, bound fractions were normalized to input fractions. Each bar represents the mean \pm standard deviation of the six independent experiments. These set of genes were then used to determine the threshold values of the microarrays.

Figure 2. Positional distribution of all identified MBD2 binding regions relative to TSSs of the nearest gene on chromosome 1. Relative distribution of MBD2 binding sites was determined through the promoter regions of chromosome 1 covered by the tiling array (-7.5 kb to +2.5 kb). Each point of the curve represents the MBD2 binding sites at 50 bp intervals.

Figure 3. The recombinant 4xMBD protein specifically shifts up probes containing methylated CpGs. Wild-type (4xMBD) or negative control (4xMBD-R22A) tetrameric MBD proteins were incubated with unmethylated (U) or methylated (M) biotin-labeled oligonucleotides that contain 5 CpG sites. The specific DNA-protein complexes were resolved on 6% polyacrylamide gels and detected by chemiluminescence as described in Materials and Methods. EBNA (Epstein-Barr Nuclear Antigen) control system, provided by

the manufacturer, was used as a positive control shift assay (Control lane). The ladder of bands observed in 4xMBD/M lane is due to complexes that contain multiple protein molecules on a single DNA molecule.

Figure 4. DNA methylation level at transcription start sites of chromosome 1. DNA methylation levels were evaluated by the normalization of methylated fragments to their CpG density at 200bp resolution and plotted according to their relative distance to TSS. This distribution indicated a strong decrease of the DNA methylation level within 1kb region surrounding TSSs.

Figure 5. Venn diagram depicting the overlap among HeLa cell genes bound by MBD2, displaying CpG methylation, and POLR2 free sites (silent genes). Of promoter sites associated with MBD2 (8257), *A*) 74% are associated with the CpG methylation mark; *B*) nearly, 69% of POLR2 negative promoters also displayed CpG methylation marks ; *C*) 65% associated with POLR2 negative promoter ; and *D*) 51% associated with both CpG methylation marks and POLR2 negative promoter. Statistical significance of overlap for all comparisons was determined to be statistically significant by χ^2 test, p values were 1.1×10^{-145} , 7.7×10^{-3} , 7.7×10^{-212} , respectively for the overlap of two parameters or by multiple correspondence analysis ($p < 0.0001$) for the overlap of the three parameters.

Figure 6. Effect of MBD2 binding distance from TSSs on transcriptional repression. The number of genes deprived of POLR2 at their TSS and exhibiting a methylated region bound by MBD2 was plotted according to the relative distance of MBD2 binding sites to TSS.

Consensus Sequence	Count	Sig Value	Coverage	CpG
ccgngcc	6066	1059.6557	42.2%	1
ccgnnccc	5194	1059.6557	40.4%	2
ccnnnnnnncgc	5169	1059.6557	39.5%	1
ccnnnnnnnnncgc	5052	1059.6557	39.4%	1
ccgnnnccc	4739	1059.6557	38.5%	1
ccnnnnnnnnncgc	4840	1059.6557	38.5%	1
ccnnnnnnnnnnncgc	4888	1059.6557	38.3%	1
ccnncgc	4857	1059.6557	38.3%	1
ccgnnnccc	4612	1059.6557	37.3%	1
ccgnnnncgc	4972	1059.6557	36.3%	2
ccnnnnnnnnncgc	4119	1059.6557	35.1%	1
ccgngcc	4232	1059.6557	34.9%	1
ccnnnnnnnnnnncgc	4191	1059.6557	34.6%	1
ccgnnncc	4170	1059.6557	34.5%	1
ccnnnnnnnnncgc	4125	1059.6557	34.3%	1
ccnnnnnnnnnnncgc	3996	1059.6557	33.8%	1
aggnnccg	3577	1059.6557	33.1%	1
ccgnncgc	4069	1059.6557	32.9%	2
ccnnnnnnncgc	3886	1059.6557	32.4%	1
ccnnnnnnncgc	3717	1059.6557	32.3%	1

Table 1. Top common motifs identified from MBD2-ChIP fragments. MBD2-binding DNA sequences were obtained from the analysis of MBD2-Immunoprecipitated chromatin fragments, on Human Promoter 1.0R Arrays (Affymetrix). Consensus sequences were identified using Scope motif finder web site (<http://genie.dartmouth.edu/scope/>). Sig value is a measure of how likely the consensus in question is to have been this over-represented. It is the negative logarithm of an expectation. Coverage indicates the percentage of the genes that contain at least one instance of the consensus sequence.

Name	p-value	Number of molecules
Gene expression	$1.81 \times 10^{-08} - 1.66 \times 10^{-02}$	469
Cell death	$1.43 \times 10^{-07} - 1.66 \times 10^{-02}$	589
Cell cycle	$2.37 \times 10^{-06} - 1.66 \times 10^{-02}$	270
Cellular assembly and organization	$2.60 \times 10^{-06} - 1.60 \times 10^{-02}$	195
DNA replication, recombination, and repair	$3.55 \times 10^{-06} - 1.25 \times 10^{-02}$	162

Table 2. Top Molecular and Cellular Functions. The silent genes exhibiting a methylated 5'end bound by MBD2 were analyzed using the Ingenuity Pathway Analysis software (<http://www.ingenuity.com/>).

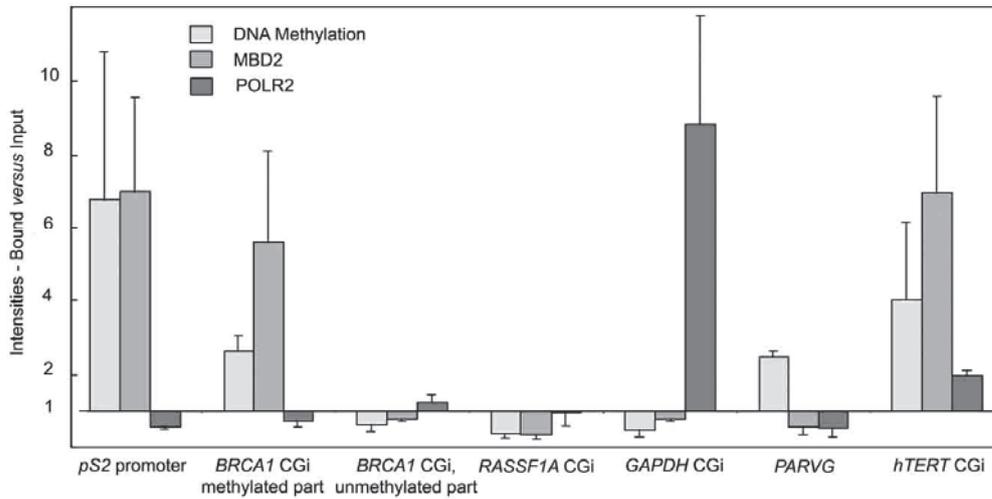


Figure 1. Efficiency and specificity of ChIP and MeDP assays

After sonication, DNAs or cross-linked chromatin were, respectively, precipitated with 4xMBD recombinant proteins in MeDP assays or with anti-MBD2 or anti-POLR2 antibodies in ChIP assays. To determine the efficiency and the specificity of the assays to use them for microarray analysis, 7 regions were analysed by quantitative PCR. These 7 regions were chosen to represent a panel of combination between methylated/unmethylated and MBD2- and POLR2- bound/unbound CGI/non-CGI regions. For each experiments and each regions, bound fractions were normalized to input fractions. Each bar represents the mean \pm standard deviation of the six independent experiments. These set of genes were then used to determine the threshold values of the microarrays.

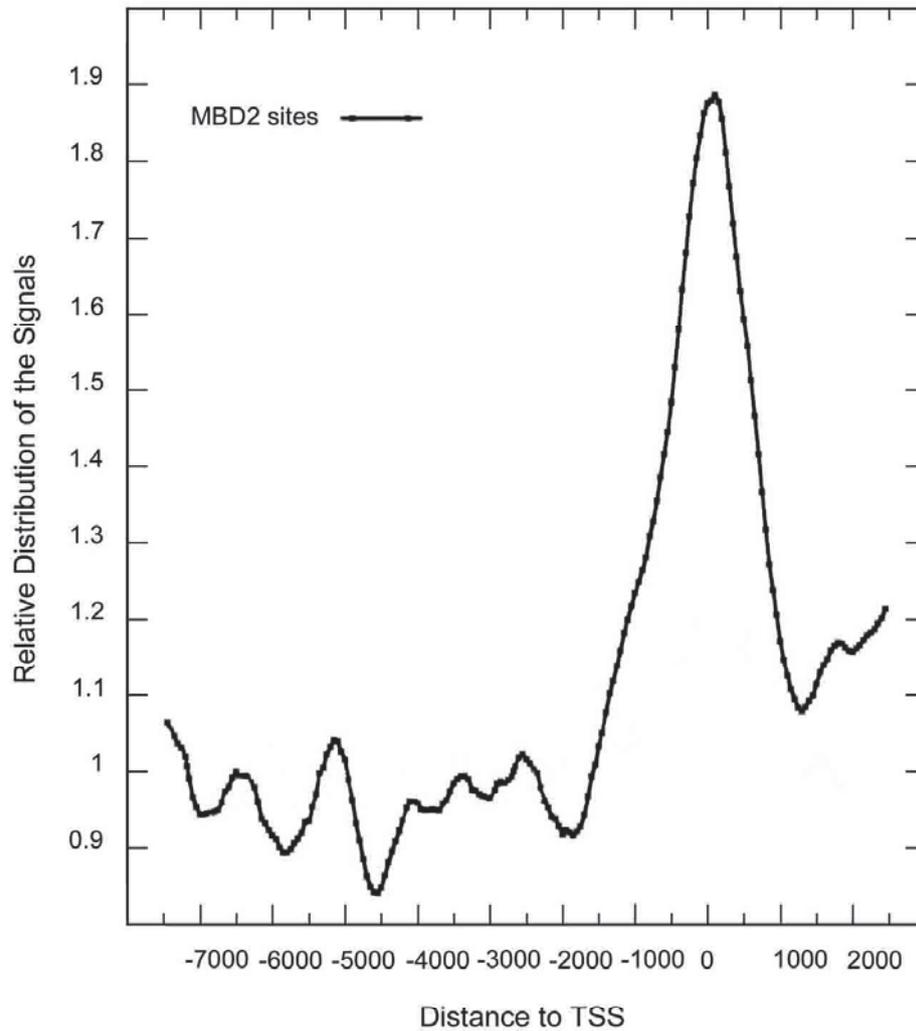


Figure 2. Positional distribution of all identified MBD2 binding regions relative to TSSs of the nearest gene on chromosome 1.

Relative distribution of MBD2 binding sites was determined through the promoter regions of chromosome 1 covered by the tiling array (-7.5 kb to +2.5 kb). Each point of the curve represents the MBD2 binding sites at 50 bp intervals.

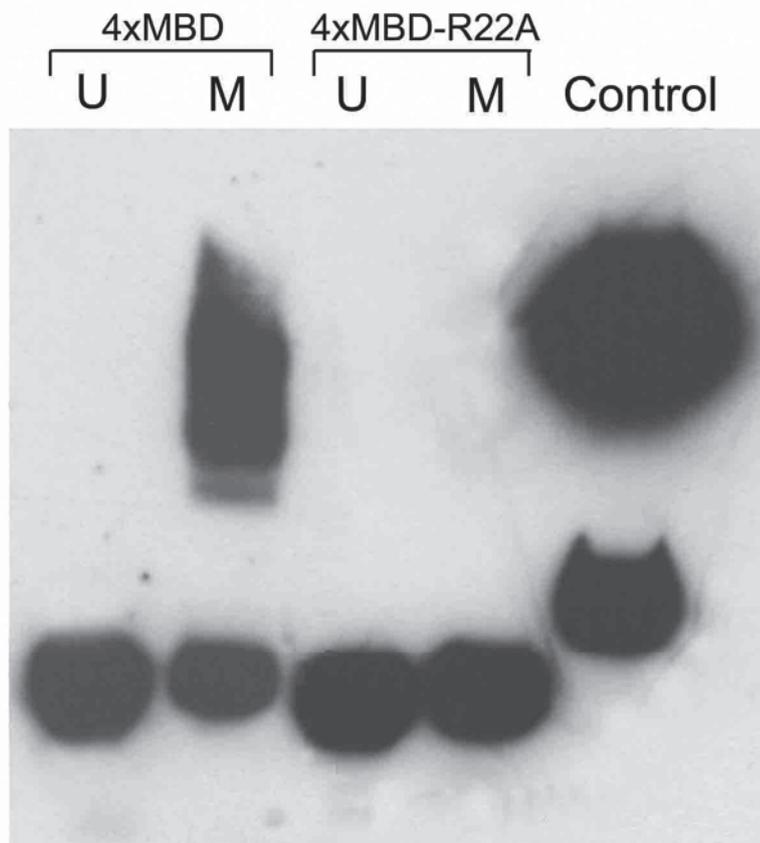


Figure 3. The recombinant 4xMBD protein specifically shifts down probes containing methylated CpGs

Wild-type (4xMBD) or negative control (4xMBD-R22A) tetrameric MBD proteins were incubated with unmethylated (U) or methylated (M) biotin-labeled oligonucleotides that contain 5 CpG sites. The specific DNA-protein complexes were resolved on 6% polyacrylamide gels and detected by chemiluminescence as described in Materials and Methods. EBNA (Epstein-Barr Nuclear Antigen) control system, provided by the manufacturer, was used as a positive control shift assay (Control lane).

The ladder of bands observed in 4xMBD/M lane is due to complexes that contain multiple protein molecules on a single DNA molecule.

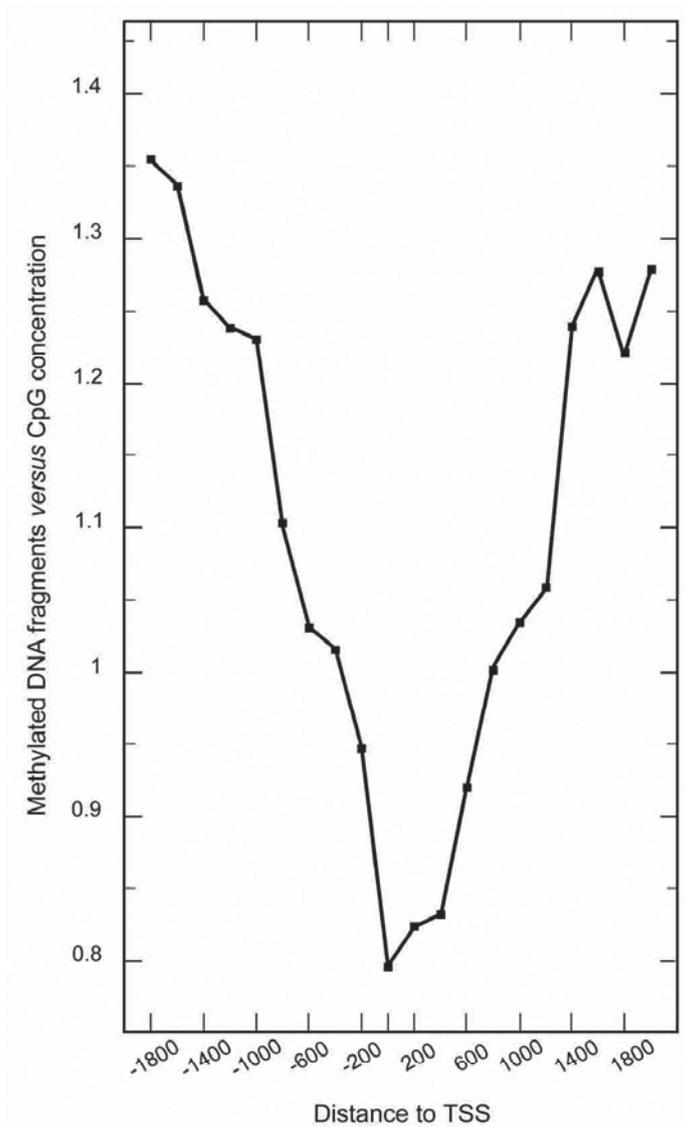


Figure 4. DNA methylation level at transcription start sites of chromosome 1
 DNA methylation levels were evaluated by the normalisation of methylated fragments to their CpG density at 200bp resolution and plotted according to their relative distance to TSS. This distribution indicated a strong decrease of the DNA methylation level within an 1kb region surrounding TSSs.

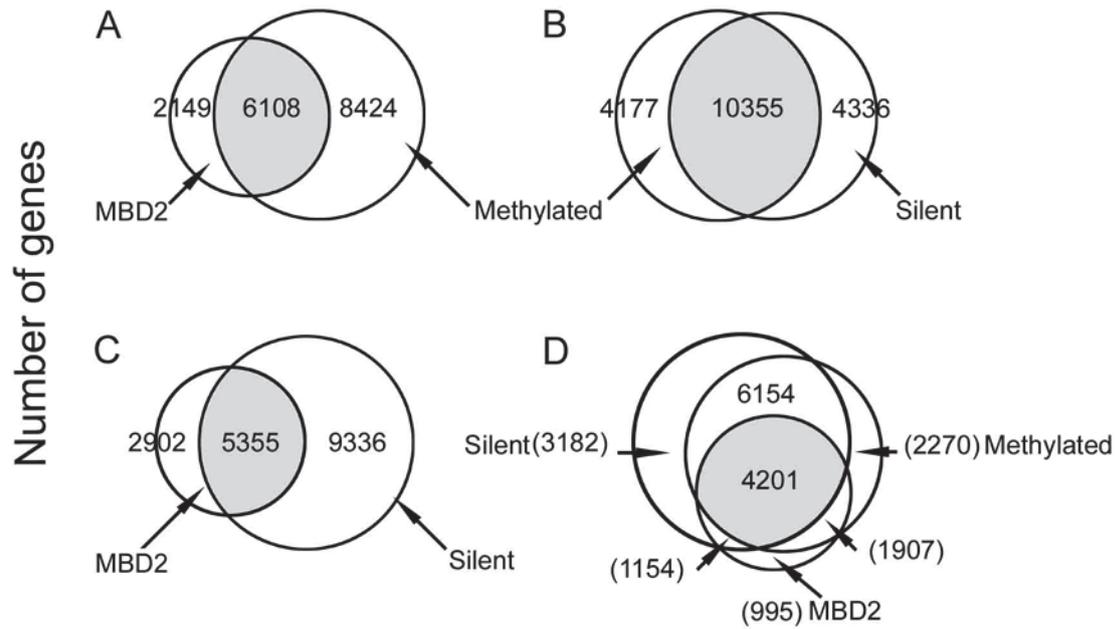


Figure 5. Venn diagram depicting the overlap among HeLa cell genes bound by MBD2, displaying CpG methylation, and POLR2 free sites (silent genes).
 Of promoter sites associated with MBD2 (8257), 74% are associated with the CpG methylation mark (A); nearly, 69% of POLR2 negative promoters also show the CpG methylation marks (B); 65% associated with POLR2 negative promoter (C); and 51% associated with both CpG methylation marks and POLR2 negative promoter (D).
 Overlap for all comparisons was determined to be statistically significant by χ^2 test, p values were 1.1×10^{-145} , 7.7×10^{-3} , 7.7×10^{-212} , respectively for the overlap of two parameters or by multiple correspondence analysis ($p < 0.0001$) for the overlap of the three parameters.

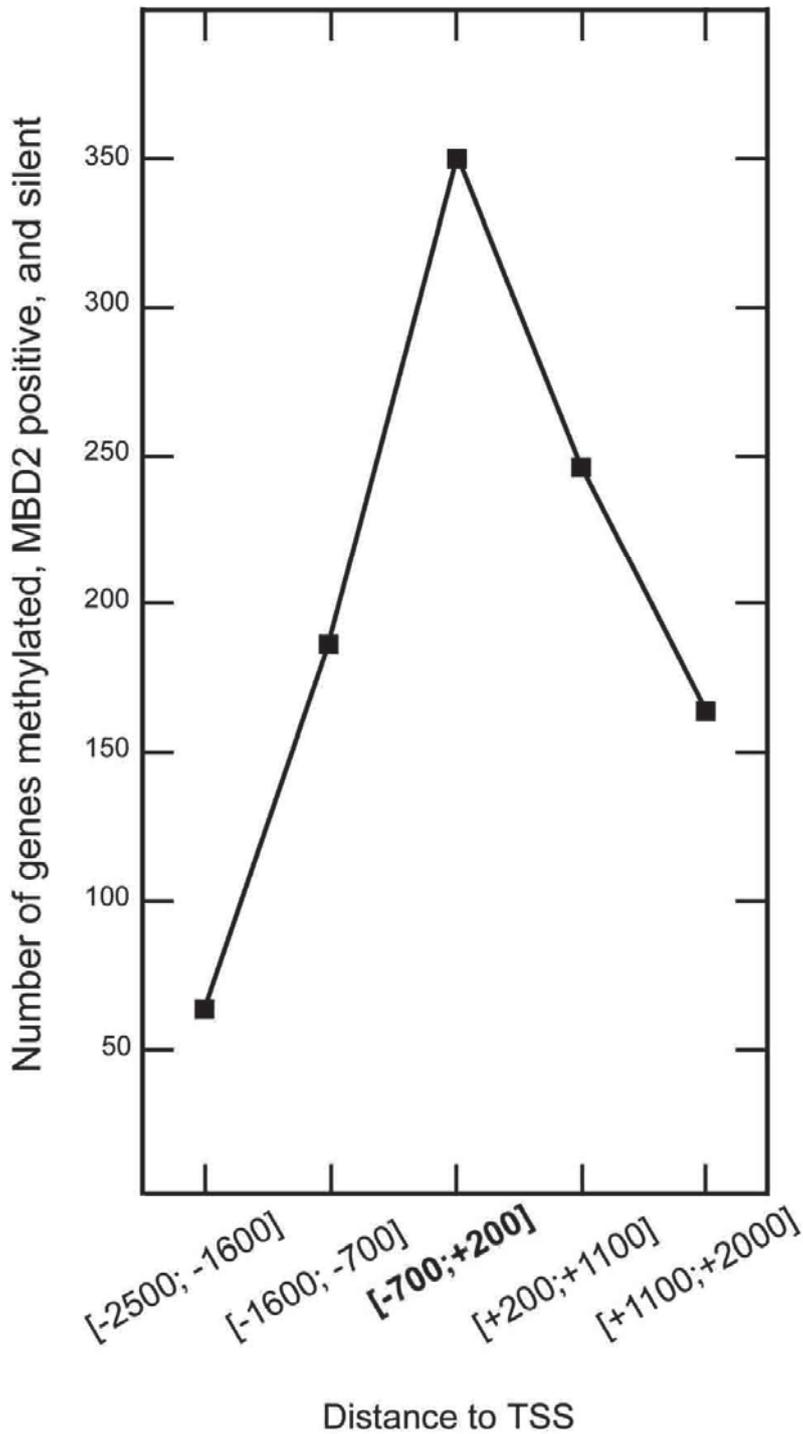


Figure 6. Effect of MBD2 binding distance from TSSs on transcriptional repression

The number of genes deprived of POLR2 at their TSS and exhibiting a methylated region bound by MBD2 was plotted according to the relative distance of MBD2 binding sites to TSS.

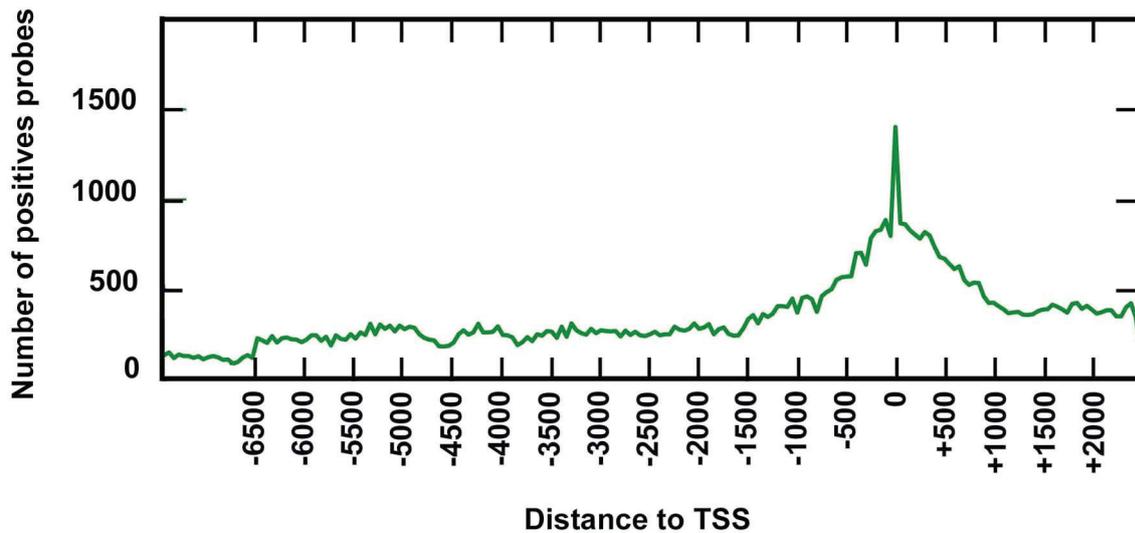


Figure S1. Positional distribution of all identified MBD2 binding regions relative to TSSs of the nearest gene on chromosome 1, raw data.

Distribution of MBD2 binding sites was determined through the promoter regions of chromosome 1 covered, at 30 bp resolution, by the tiling array (-7.5 kb to +2.5 kb). Each point of the curve represents the number of positive probes.

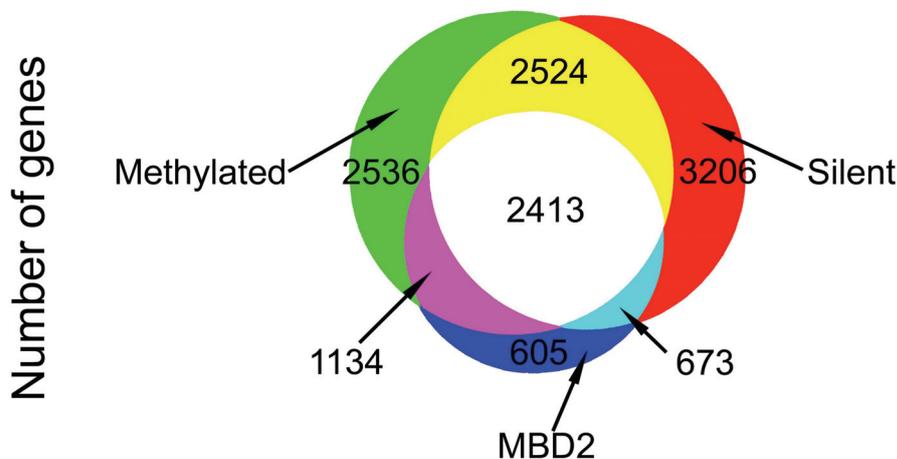


Figure S2. Venn diagram depicting the overlap among HeLa cell genes bound by MBD2, displaying CpG methylation, and POLR2 free sites (silent genes), from a filtered promoter list.

Among the 23,054 putative promoters, analyzed in this study, 12,486 have been annotated with high-confidence using three different data banks. Of promoter sites associated with MBD2 (4825), 74% are associated with CpG methylation marks (A); 64% associated with POLR2 negative promoters (B); 59% of POLR2 negative promoters also show CpG methylation marks (C); and 50% associated with both CpG methylation marks and POLR2 negative promoters.

Gene name	ChIP	ChIP on chip	Validation
ESR1	+	+	V
p16	+	+	V
RASSF1A	-	-	V
SYK	-	-	V
CDH1	-	-	V
TMS1	-	-	V
DAPK1	+	+	V
p73	-	-	V
RIZ	-	-	V
GSTP1	-	-	V
MLH1	-	-	V
CRABP1	-	-	V
RARB	-	-	V
p14	+	+	V
CDH13	+	-	NV
PR	-	+	NV
MGMT	-	+	NV
p15	-	+	NV

Table S1. List of MBD2-associated genes identified either by direct ChIP experiments (1) or by ChIP on chip analysis.

Genes exhibiting MBD2-binding sites are indicated by “+”, and absence of MBD2 by “-“. ChIP, identification of MBD2-binding sites by direct ChIP experiments; Chip on chip, data obtained from our own ChIP on chip experiments. V, indicates concordance; NV, indicates discordance between both methods.

1. Lopez-Serra, L., Ballestar, E., Fraga, M.F., Alaminos, M., Setien, F. and Esteller, M. (2006) A profile of methyl-CpG binding domain protein occupancy of hypermethylated promoter CpG islands of tumor suppressor genes in human cancer. *Cancer Res*, **66**, 8342-8346.

Gene name	MLM	MeDP	<i>Cfol</i> site	Validation
CTCF	+	+	-950 & -927	V
GATA5	-	-	-740	V
WNT2	+	+	-69	V
ADAMTS18	+	+	-63	V
DKK3	+	+	-608	V
APC	-	-	-59	V
SFRP1	+	+	-58 & -56	V
TP73	+	+	-51 & -36	V
GpX3	+	+	50	V
SFRP2	+	+	-487	V
MYOD1	+	+	-413	V
TERT	+	+	-367	V
NELL1	+	+	-312 & -284	V
CDH1	+	+	-227	V
SFRP5	+	+	-174	V
CDH13	+	+	-16	V
RASSF1A	-	-	-135	V
DAPK1	-	-	-1090	V
MGMT	+	+	+152	V
CDKN2A	-	-	-418	V
SOCS3	+	+	147	V
MLH1	-	+	-636	N V
DKK2	-	+	-592	N V
SMAD4	-	+	-36, -34 & -32	N V
SFRP4	+	-	-274	N V
RPRM	+	-	-236	N V
DKK1	+	-	-155	N V

Table S2. List of genes exhibiting a methylated sequence identified either by enzymatic digestion or by MeDP.

MLM, methylation status of 5' end of genes determined by a new approach termed Methylation Ligation-dependent Macroarray (MLM, a method based on a Multiplex Ligation-dependent Probe Amplification protocol MLPA), using *Cfol*, a methylation sensitive restriction enzyme; MeDP, methylation status determined from our own MeDP on chip experiments. Methylated genes are indicated by "+" and unmethylated by "-". V, indicates concordance; NV, indicates discordance between both methods.

Article 4

*Specific Distribution of the Methyl-CpG Binding Domain Protein 2
at methylated DNA sequences during tumoral progression, in a
syngenic model of breast cancer*

Laury Perriaud, Anne-Pierre Morel, Joel Lachuer and Robert Dante

En cours de rédaction

Contexte d'étude

Dans les cellules cancéreuses, des profils de méthylation de l'ADN aberrants ont été mis en évidence depuis plus de 20 ans. En effet, une hypométhylation globale du génome entraînant l'expression ectopique de certains gènes et une instabilité génomique est associée à une hyperméthylation de certains îlots de CpG, induisant la répression de certains gènes suppresseurs de tumeurs. De plus, il est maintenant établi que les doublets CpG méthylés peuvent être reconnus par différentes familles de protéines liant l'ADN méthylé (MBP). Aussi, l'altération des profils de méthylation de l'ADN pourrait conduire à une modification des profils de liaison de ces protéines et participer aux changements d'expression génique dans les cancers. Le but de notre étude est donc de comprendre quand et comment les MBP et plus particulièrement la protéine MBD2 ségrégent entre les sites méthylés au cours de la cancérogenèse, et comment ces événements traduisent le signal de méthylation en une altération conférant un avantage sélectif aux cellules cancéreuses.

Nous avons privilégié la protéine MBD2 comme modèle car elle semble, au moins pour certains cancers, impliquée dans le processus tumoral. En effet, l'inhibition de *Mbd2*, par invalidation chez la souris *Apc^{Min/+}* ou par injections d'ARN antisens ciblant l'ARNm *Mbd2* chez les souris immunodéficientes et greffées avec des tumeurs humaines, a un fort effet anti-tumoral. De plus, en collaboration avec l'équipe de M. Esteller (CNIO à Madrid) et au cours de l'article 3, nous avons montré que cette protéine occupe une place importante de répresseur dépendant de la méthylation, au moins dans les cellules cancéreuses HeLa.

Stratégie d'étude

Notre objectif est de suivre les modifications des profils de méthylation de l'ADN et d'identifier les conséquences de ces variations sur la redistribution de la protéine MBD2 et la régulation génique au cours de la progression tumorale. Ces recherches s'inscrivent dans l'étude des liaisons des protéines à MBD aux promoteurs, jusqu'ici principalement menées dans des lignées cellulaires cancéreuses. Or, dans ces travaux, il manque deux témoins : la contre-partie saine et les différentes étapes de la progression tumorale, en particulier les événements menant à l'immortalisation. Aussi, nous avons utilisé un modèle fondé sur celui établi par l'équipe de Weinberg (Elenbaas, Spirio et al. 2001) et recomposé dans notre

laboratoire (Morel, Lievre et al. 2008). Ce modèle utilise des cellules épithéliales mammaires humaines saines (HMEC) transfectées avec la transcriptase inverse de la télomérase humaine (cellules HMEC-*hTERT*) pour l'acquisition de l'immortalisation, les antigènes t et T de SV40 (cellules HMLE) puis avec différents oncogènes dont *Ras* humain pour leur transformation (cellules HMLER) (**Figure 1**). Aussi, en utilisant ces lignées, nous avons suivi la distribution de la protéine MBD2 dans un modèle homogène reproduisant la progression tumorale mammaire.

De plus, avec l'étude réalisée précédemment (article 3), j'ai mis en place la stratégie et les outils nécessaires pour étudier de façon globale les sites de méthylation, les sites de fixation de la protéine MBD2 et les sites ne présentant pas d'activité transcriptionnelle. A ces études globales réalisées par des approches « MeDP-on-chip » et « ChIP-on-chip », j'ai rajouté celle de l'acétylation de l'histone H3, une marque de décondensation de la chromatine associée à une activité transcriptionnelle.

Matériels et méthodes

Culture cellulaire

Les cellules épithéliales mammaires humaines (HMEC) proviennent de Lonza (Lonza, Vervier, Belgium) et ont été cultivées selon leurs recommandations, en présence de 5% de CO₂, à 37°C en chambre humide. Leurs dérivés (HMEC-*hTERT*, HMLE, fournis par R.Weinberg et HMLER générées par nos soins par transfection virale (Morel, Lievre et al. 2008)) ont été cultivées dans un mélange de milieu de Eagle modifié par Dulbecco (DMEM) et Ham's F-12 (1 :1) avec L-alanyl-L-Glutamine (GIBCO™ DMEM/F-12 L-GlutaMAX™ I) (Invitrogen, Carlsbad, CA), supplémenté avec 10% de veau fœtal (Lonza), de l'insuline humaine à 10µg/ml (Actrapid 100UI/ml, Novo Nordisk Pharmaceutique SAS), de l'EGF humaine à 10ng/ml (Promocell, Heidelberg, Germany), de l'hydrocortisone à 0,5µg/ml (Sigma, L'isle d'Abeau, France) et 1% d'antibiotiques pénicilline et streptomycine (Sigma) et pour HMLER, 0,5µg/ml de puromycine sont ajoutées pour leur sélection (Invivogen, Toulouse, France).

Immunoprécipitations de chromatine (ChIP)

Ce protocole a été décrit dans les articles 1, 2 et 3. Les complexes nucléoprotéiques issus des cellules HMEC-*hTERT* ou HMLER ont été précipités avec 20µl d'anticorps polyclonaux dirigés contre la protéine MBD2 produits à notre demande par Covalab et validés au cours de l'article 2 ou d'anticorps polyclonaux anti-ARN polymérase II (8WG16, Covance, California) ou d'anticorps polyclonaux anti-histones H3 acétylées (Millipore, Billerica, MA). Comme aucun gène candidat n'a été étudié dans ce modèle pour ces paramètres, la première validation des résultats des ChIP a consisté à réaliser des PCR quantitatives à partir d'échantillons indépendants (n=3) pour différentes régions promotrices (**Tableau 1**) et a vérifié la similitude de leurs profils d'amplification entre échantillons.

Des immunoprécipitations de chromatine ont également été effectuées à partir de 20µl d'anticorps monoclonaux anti-Kaiso (6F/6F8, Abcam, Paris, France) ou de 4µl de sérum de lapin immunisé, fourni par P.A. Defossez (Paris, France). Leurs validations, reposant sur le même principe que précédemment, sont en cours d'analyse.

Précipitation d'ADN méthylé (MeDP)

Les ADN issus des cellules HMEC, HMEC-*hTERT* et HMLER ont été précipités selon le principe défini au cours de l'article 3 à partir de nouveaux lots vérifiés de la protéine 4xMBD. Pour tester l'efficacité des immunoprécipitations d'ADN méthylé, des PCR quantitatives ont été réalisées à partir de différentes régions promotrices, la région « ilot *BRCA1* » servant de contrôle positif (enrichissement) (**Tableau 1**).

« MeDP-on-chip » et « ChIP-on-chip »

Les techniques de « MeDP-on-chip » et « ChIP-on-chip » et leur méthode d'analyses sur puces promoteurs ont été décrites au cours de l'article 3.

La validation des données issues des puces promoteurs sera complétée par l'amplification de nouvelles régions promotrices.

Quantifications des transcrits MBD2, MBD1 et MeCP2

Le taux des transcrits *MBD2* a été quantifié par RT-PCR compétitive, celui des transcrits *MBD1* et *MeCP2* normalisé par rapport au gène de ménage *PGK1* après PCR quantitative, comme décrit au cours de l'article 2.

Western Blot

Le protocole a été décrit dans l'article 2. Les membranes de nitrocellulose ont été saturées par du TBS-Tween 0,5%+ 5% lait ou 5% d'albumine sérique bovine (BSA). Les anticorps primaires anti-Kaiso (Abcam) ont été dilués au 1/500^e dans du TBST+5% BSA alors que les anticorps anti-ZBTB4 (sérum fourni par P.A Defossez) et anti-Ku80 (BD Biosciences, Le Pont de Claix, France), au 1/3000^e dans du TBST+5% lait. Les anticorps secondaires couplés à l'HRP sont soit des anticorps de lapin anti-Ig de souris (P0260, Dakocytomation, Trappes, France) dilués au 1/3000^e dans du TSBT+5% lait soit des anticorps de mouton anti-Ig de lapin dilués au 1/3000^e dans du TBST+5% BSA (P0448, Dakocytomation).

Les tailles attendues sont de 100kDa pour Kaiso, 150 kDa pour ZBTB4 et 80 kDa pour Ku80.

Résultats

Validations des MeDP, CHIP et de leurs hybridations sur puces promoteurs

Pour cartographier les profils de méthylation de l'ADN au cours de la progression tumorale, des immunoprécipitations d'ADN (MeDP) de cellules HMEC, HMEC-*hTERT* et HMLER avec la protéine recombinante 4xMBD ont été effectuées afin de sélectionner l'ADN méthylé. La spécificité de ces immunoprécipitations a été évaluée par comparaison de l'enrichissement en ADN méthylé des fractions précipitées ou « bound » par rapport aux fractions initiales ou « input » en analysant par PCR quantitatives des régions promotrices. Comme aucune donnée dans la littérature n'est disponible par rapport aux profils de méthylation de gènes candidats dans ce modèle de progression tumorale, nous avons réalisé différentes MeDP à partir d'échantillons indépendants et vérifié la similitude des profils d'amplification des diverses régions promotrices. Comme la région « ilot *BRCA1* » a été décrite méthylé et fixée par la protéine MBD2 dans de nombreuses lignées cellulaires et que la région amplifiée pour *CDH13* ne compte pas de CpG, nous avons utilisé ces régions respectivement comme témoins d'enrichissement et d'appauvrissement des fractions précipitées en ADN méthylé. Les régions promotrices d'*EIF3A* (aussi appelée *EIF3S10*), *PRPH2*, *DAPK1* et ilot *RASSF1A* présentent un panel de régions avec des densités de CpG différentes. Ainsi, il apparaît que dans les cellules HMEC, la région « ilot *BRCA1* » est méthylée alors que les régions promotrices d'*EIF3A*, *CDH13*, *PRPH2*, *DAPK1* et l'ilot

RASSF1A sont non méthylés (**Figure 2A**). Dans les cellules HMEC-*hTERT* et HMLER, les profils restent les mêmes à l'exception de *PRPH2* qui devient méthylé dès HMEC-*hTERT* et de l'îlot *RASSF1A* qui se méthyle progressivement au cours de la progression tumorale (**Figures 2B et 2C**).

Après ces validations, les précipitations d'ADN par la protéine 4xMBD ont été ensuite hybridées sur des puces promoteurs « Affymetrix Human promoter 1.0R Arrays » par nos collaborateurs de la plateforme ProfileXpert (Lyon), comme décrit au cours de l'article 3. Puis, nous avons effectué l'analyse des données grâce aux logiciels GCOS, TAS, IGB et CisGenome. Les données issues des validations de MeDP ont servi à poser les seuils, à savoir $\log_2 > 0,1$ pour la méthylation pour les différentes lignées cellulaires (**Tableau 2**). Une seconde validation des puces promoteurs est en cours d'analyse, avec amplification par PCRq de nouvelles régions promotrices (résultats non représentés).

En parallèle, des ChIP ont été réalisées avec des anticorps dirigés contre la protéine MBD2 ou contre l'ARN polymérase II ou contre les histones H3 acétylées, à partir de chromatine de cellules HMEC-*hTERT* et HMLER. La lignée cellulaire HMEC n'a pas été étudiée car sa prolifération est très lente et donc inadaptée pour les ChIP effectués ainsi que pour le temps dont je disposais. La stratégie de validation et de fixation des seuils des puces reste la même que celle employée pour les MeDP. Dans les deux lignées, la protéine MBD2 est présente sur la région « îlot *BRCA1* » comme attendu, est absente sur les régions non méthylées ainsi que sur la région promotrice de *PRPH2* malgré son hyperméthylation, et suit le profil de méthylation de l'îlot *RASSF1A* qui devient une cible de la protéine MBD2 (**Figure 2**). Concernant l'absence de fixation de l'ARN polymérase II et par conséquent l'inhibition transcriptionnelle, il apparaît que tous les gènes étudiés sont silencieux dans HMEC-*hTERT* alors que pour HMLER, *EIF3A* et *CDH13* présentent respectivement une forte et faible présence de l'ARN polymérase II sur leurs promoteurs, ce qui est concordant avec l'absence de méthylation de l'ADN et de la protéine MBD2. Enfin, il semble que la présence d'histones H3 acétylées ne soit détectée que pour la région promotrice du gène *EIF3A* dans les cellules HMLER, concordant avec le profil observé pour l'ARN polymérase II. Aussi, **toutes les données des MeDP et ChIP obtenues pour les 6 régions promotrices étudiées sont validées pour chaque paramètre par la répétition des profils observés dans différents échantillons et sont concordantes entre elles (Figure 2)**. Puis, de la même façon que pour les MeDP, les données issues des validations de ChIP ont servi à poser les seuils, à savoir $\log_2 > 0,5$ pour la présence de MBD2, $\log_2 < -0,1$ pour l'absence d'ARN polymérase II, et

$\log_2 > 0,4$ et $\log_2 > 0,3$ respectivement pour HMEC-*hTERT* et HMLER concernant la présence d'histones H3 acétylés (**Tableau 2**). Une seconde validation de ces données vient d'être commencée (résultats non représentés).

Altération des profils de méthylation de l'ADN au cours de la progression tumorale

Il a été décrit que les cellules cancéreuses présentent un profil de méthylation altéré par rapport aux cellules normales. Aussi, nous avons réalisé une cartographie des sites méthylés par « MeDP-on-chip » en comparant par hybridation sur puces promoteurs les fractions précipitées avec la protéine 4xMBD aux fractions initiales à différentes étapes de la progression tumorale. L'analyse des données révèle l'existence de 8416 promoteurs (région comprise entre -700 et +200pb autour du site d'initiation de la transcription) présentant au moins un site méthylé pour les cellules HMEC, 8465 pour les cellules HMEC-*hTERT* et 8104 pour les cellules HMLER (**Tableau 2**), ce qui représente une **diminution de 4% entre les stades précoces et tardif de la cancérogenèse**. Le recoupement de ces données à l'aide d'Access (logiciel Microsoft) a ensuite permis de déterminer la redistribution de la méthylation de l'ADN au cours de la cancérogenèse. Ainsi, les cellules HMEC partageraient 6494 (77%) de leurs sites méthylés avec HMEC-*hTERT* et 6202 (74%) avec HMLER. Les 3 lignées cellulaires auraient en commun 5119 sites méthylés, signifiant que ces sites sont méthylés à tous les stades de la progression tumorale. Enfin, 839 sites sont uniquement méthylés dans HMEC, 842 dans HMEC-*hTERT* et 773 dans HMLER, ce qui représente pour chaque lignée 10% des sites totaux méthylés (**Figure 3**). Toutes ces données suggèrent qu'il **existe une redistribution modérée de la méthylation de l'ADN au cours de la progression tumorale**.

De plus, nous avons voulu classer les promoteurs méthylés en fonction de leur fréquence en $CpG_{\text{observés/théoriques}}$ et de leur ratio C+G afin de préciser quelle(s) catégorie(s) de promoteurs étai(en)t principalement concernée(s) par ces altérations de méthylation. Pour cela, nous avons utilisé la liste générée par Weber et ses collaborateurs qui ont classé les promoteurs en 3 catégories selon les critères suivants : les promoteurs riches en CpG (high-CpG promoters ou HCP) correspondent à des régions de 500 pb avec une densité supérieure à 0.75 et un taux de C+G supérieur à 55%, les promoteurs pauvres en CpG (low-CpG

promoteurs ou LCP) ont une densité inférieure à 0.48 et les promoteurs intermédiaires (intermediate CpG promoteurs ou ICP) ne sont ni des HCP ni des LCP. La distribution des promoteurs sur la puce est de 62% de HCP, 13% d'ICP et 25% de LCP. Pour les sites méthylés communs, on observe une distribution différente avec 81% de HCP, 9% LCP et 10% ICP. Ceci suppose donc que **les principales cibles de l'altération de la méthylation de l'ADN au cours de la progression tumorale sont les LCP et dans une plus faible proportion les ICP** alors que **les HCP semblent plutôt exemptés**. Les sites spécifiques de chaque lignée ont une répartition similaire entre HCP et LCP (de l'ordre de 40%) et de 16% pour les ICP (**Figure 3**), indiquant encore une fois que la méthylation de l'ADN paraît plus susceptible d'être conservée au sein des régions riches en CpG.

La protéine MBD2 garde un rôle répressif dépendant de la méthylation tout au long de la progression tumorale

Nous avons vu que la méthylation de l'ADN est une marque épigénétique qui n'a pas d'impact en elle-même mais qui intervient notamment dans la répression transcriptionnelle en empêchant la fixation de facteurs de transcription nucléaires ou plus majoritairement en étant « lue et traduite » par des protéines liant l'ADN méthylé recrutant des complexes enzymatiques remodelant la chromatine afin de la rendre transcriptionnellement inactive. De plus, nous avons décrit au cours de l'article 3 que la protéine MBD2 est un vrai répresseur dépendant de la méthylation de l'ADN dans les cellules cancéreuses HeLa.

Aussi, afin de suivre la distribution de la protéine MBD2 tout au long de la progression tumorale, nous avons réalisé en plus de l'étude des profils de méthylation par « MeDP-on-chip », une cartographie par « ChIP-on-chip » des sites de fixation de la protéine MBD2, ainsi que des sites ne présentant pas d'ARN polymérase II donc sans activité transcriptionnelle et des sites acétylés sur les histones H3 marquant une décondensation de la chromatine favorable à une activation transcriptionnelle dans les cellules HMLER (cellules cancéreuses du modèle) et des cellules HMEC-*hTERT* (cellules immortalisées représentant un stade précoce dans la progression tumorale).

Dans les **cellules HMEC-*hTERT***, la majorité des sites fixés par la protéine MBD2 (~72%) présentent un profil méthylé (**Figure 4A**), avec une corrélation statistiquement

significative (test χ^2 : $p= 1,98 \times 10^{-8}$). Cette protéine paraît donc avoir une activité **majoritairement méthylation-dépendante et se lier de façon spécifique**. A l'inverse, près de 32% des sites méthylés apparaissent liés par la protéine MBD2 (**Figure 4A**), suggérant un **rôle prépondérant de cette protéine parmi celles liant l'ADN méthylé**.

Les résultats montrent également que 76% des sites méthylés sont silencieux et que 72% des sites réprimés sont méthylés (**Figure 4B**), marquant une **corrélation inverse statistiquement significative** (test χ^2 : $p= 4,09 \times 10^{-42}$) **entre la présence de méthylation et l'expression génique et plaçant la méthylation comme élément important des mécanismes répressifs**. De plus, 66% des sites fixés par la protéine MBD2 sont associés à une inhibition transcriptionnelle de façon statistiquement significative (test χ^2 : $p= 2 \times 10^{-25}$), suggérant que cette protéine a un rôle **répressif majoritaire et spécifique (Figure 4C)**.

Il apparaît également que 42% des sites fixés par MBD2 sont méthylés et silencieux et ~30% des gènes méthylés et silencieux sont fixés par la protéine MBD2 (**Figure 4D**), lui conférant un **important et spécifique rôle de répresseur dépendant de la méthylation**. Enfin, le croisement des 1908 sites méthylés, fixés par la protéine MBD2 et associés à la répression transcriptionnelle avec la liste des sites acétylés en H3 montre que 72% des 1908 sites ne sont pas acétylés en H3. On pourrait ainsi en déduire que **la répression initiée par la fixation de la protéine MBD2 sur des sites méthylés est associée à une chromatine condensée**.

Dans les **cellules HMLER, les résultats sont similaires**. En effet, les sites fixés par la protéine MBD2 sont méthylés à 71% (test χ^2 : $p = 9,46 \times 10^{-22}$), associés à une répression transcriptionnelle à 59% (test χ^2 : $p= 1,31 \times 10^{-36}$). De plus, 45% des promoteurs liés par MBD2 sont méthylés et silencieux et 29% des gènes méthylés et silencieux sont fixés par la protéine MBD2 (résultats non représentés). La répression transcriptionnelle initiée par la protéine MBD2 est associée à 76% à des histones H3 non acétylés. Ces données suggèrent donc que **la répression initiée par la protéine MBD2 est dépendante de la méthylation de l'ADN et associée à une compaction chromatinienne aussi dans les cellules HMLER. La protéine MBD2 garderait donc le même rôle tout au long de la progression tumorale**. De plus, bien que ce rôle soit plus marqué dans les cellules HeLa (51% des sites liés par MBD2 sont méthylés et silencieux et ~41% des gènes méthylés et silencieux sont fixés par la protéine MBD2) que dans les cellules HMLER, **cette fonction ne semble pas spécifique d'un type cellulaire**.

La protéine MBD2 est redistribuée au cours de la progression tumorale

Mise en évidence de la redistribution de la protéine MBD2

Nous avons vu au cours de l'introduction, dans le chapitre 2, que les sites liés par la protéine MBD2 ont été principalement identifiés dans des lignées cellulaires cancéreuses et que très peu de données concernant les cellules normales sont disponibles. D'autre part, l'évolution des profils de liaisons de la protéine MBD2 au cours de la cancérogenèse n'a pas été étudiée.

Aussi, nous avons voulu réaliser une comparaison des sites présentant la protéine MBD2 dans les cellules HMEC-*hTERT* et HMLER afin de suivre à différents stades de la progression tumorale la distribution de cette protéine. Ainsi, l'analyse des données « ChIP anti-MBD2-on-chip » révèle que 3826 promoteurs dans les cellules HMEC-*hTERT* et 3736 dans les cellules HMLER présentent au moins une région fixée par la protéine MBD2, montrant un **nombre similaire de sites (Tableau 2)**. Par contre, **près de la moitié de ces sites ne sont pas communs entre ces deux lignées cellulaires (Figure 5)**, montrant une **forte redistribution de la protéine MBD2 au cours de la progression tumorale**.

La redistribution de la protéine MBD2 n'est pas associée à un effet de concentration

Nous avons voulu comprendre comment ces changements de position intervenaient. Comme il a été montré que l'abondance de la protéine pouvait jouer un rôle dans la distribution de la protéine MBD2, nous avons réalisé une quantification des transcrits *MBD2* par RT-PCR compétitive à partir de trois échantillons d'ARN total indépendants à différentes étapes de la progression tumorale, étant donné que le taux protéique MBD2 est corrélé au taux d'ARNm *MBD2*. Ainsi, les cellules HMEC présentent en moyenne $11,8 \cdot 10^6$ molécules d'ARNm *MBD2* par microgramme d'ARN total, contre $17 \cdot 10^6$ pour les cellules HMEC-*hTERT* et $23,9 \cdot 10^6$ pour les cellules HMLER (**Figure 6**). Si ces données suggèrent que le taux d'ARNm *MBD2* augmente au cours de la progression tumorale notamment entre HMEC et HMLER (test t : $p=0,005$), aucune différence significative n'est observée entre HMEC-*hTERT* et HMLER (test t : $p= 0,052$). Aussi, ces résultats laissent supposer que la **redistribution de la protéine MBD2 au cours de la progression tumorale n'est pas due à un changement de concentration protéique**.

La redistribution de la protéine MBD2 n'est pas due aux altérations de la méthylation de l'ADN

Comme la protéine MBD2 se fixe préférentiellement sur des sites méthylés, la redistribution de la protéine MBD2 pourrait simplement être le reflet des changements de profils de méthylation. Aussi, nous avons comparé les sites de méthylation entre les cellules HMEC-*hTERT* et HMLER ainsi que les sites méthylés fixés par la protéine MBD2. Il apparaît que ces lignées partagent ~75% de leurs sites méthylés et ~41% des sites méthylés fixés par la protéine MBD2 (**Figure 5**). Ceci suggère donc que **la redistribution de la méthylation a un faible impact sur la redistribution de la protéine MBD2**.

Toutes ces données suggèrent que la **redistribution de la protéine MBD2 au cours de la progression tumorale** est dépendante d'autres facteurs stabilisant la liaison de la protéine MBD2 à l'ADN méthylé ou participant à son recrutement.

Répercussions de cette redistribution sur la répression initiée par la méthylation

Nous avons vu que la protéine MBD2 semble jouer un rôle important dans la répression. Aussi, nous avons effectué un premier croisement des sites associés à la répression transcriptionnelle entre HMEC-*hTERT* et HMLER et montré plus de 75% de sites en commun (**Figure 5**), marquant une redistribution plus limitée que celle de la protéine MBD2. Ces résultats suggèrent que **la redistribution de la protéine MBD2 influence faiblement la redistribution de la répression transcriptionnelle**.

Pour mettre en évidence l'influence directe de la protéine MBD2 en tant que répresseur majoritairement dépendant de la méthylation, nous avons ensuite comparé les sites méthylés, fixés par la protéine MBD2 et associés à la répression transcriptionnelle entre les cellules HMEC-*hTERT* et HMLER et ainsi déterminé 524 sites communs alors que 1384 sites sont spécifiques de la lignée cellulaire HMEC-*hTERT* et 1153 de la lignée HMLER (**Figure 5**). Aussi, ces **cellules partagent peu de gènes réprimés suite à la fixation de MBD2 sur leurs promoteurs méthylés**.

Parmi les gènes présentant des promoteurs méthylés, liés par la protéine MBD2 et associés à la répression transcriptionnelle dans les cellules HMEC-*hTERT* et HMLER, 86% (451) présentent également un promoteur méthylé dans les cellules HMEC. Ces **451 gènes**

sont donc constamment méthylés, quelque soit le stade de la progression tumorale. On peut alors émettre l'hypothèse que cette **méthylation participe à la répression de ces gènes, nécessaire à l'établissement du phénotype mammaire.** Il faudrait déterminer le statut de fixation de la protéine MBD2 et de la répression transcriptionnelle de ces 451 gènes pour confirmer ou infirmer cette idée.

D'autre part, pour déterminer les fonctions biologiques majeures des protéines codées par ces gènes, la liste des 451 gènes a été soumise sur le site internet GOTM (Gene Ontology Tree machine : <http://bioinfo.vanderbilt.edu/gotm/>). Il semble que ces protéines participent essentiellement à la **régulation transcriptionnelle avec des propriétés de liaison à l'ADN et aux protéines (Tableau 3).**

De plus, nous avons voulu déterminer les critères associés à la distribution spécifique des autres sites dans ces deux lignées cellulaires. Nous avons alors comparé la liste des 1384 sites méthylés, liés par MBD2 et associés à la répression transcriptionnelle spécifiquement dans HMEC-*hTERT* avec les listes des sites méthylés ou les sites fixés par la protéine MBD2 ou les sites présentant une absence de l'ARN polymérase II dans HMLER. Il apparaît ainsi parmi ces 1384 sites, 40% soit 559 gènes sont toujours réprimés et présentent un promoteur méthylé alors qu'il y a perte de la fixation de la protéine MBD2 entre HMEC-*hTERT* et HMLER. La même comparaison à partir de la liste spécifique multifactorielle de HMLER et des autres listes de HMEC-*hTERT* montre que parmi les 1153 gènes, 52% soit 600 gènes sont toujours réprimés et présentent un promoteur méthylé alors qu'ils gagnent une liaison de MBD2 entre HMEC-*hTERT* et HMLER. Aussi, **certains gènes restent réprimés et possèdent toujours un promoteur méthylé entre les cellules immortalisées et les cellules cancéreuses alors que leur statut concernant la présence de la protéine MBD2 change (gain ou perte).**

Remplacement de la protéine MBD2 par une autre protéine liant l'ADN méthylé ?

La constante répression de certains gènes via la méthylation, quelque soit leur statut vis-à-vis de la fixation de la protéine MBD2, **suggère qu'il y a un remplacement mutuel entre la protéine MBD2 et un ou d'autres répresseurs dépendant de la méthylation** c'est-

à-dire par d'autres protéines liant l'ADN méthylé, afin de conserver l'inhibition transcriptionnelle de ces gènes quelque soit le stade de progression tumorale.

Aussi, je me suis intéressée au potentiel remplacement de la protéine MBD2 dans les cellules HMLER comparé aux cellules HMEC-*hTERT*. Les protéines qui pourraient remplir ce rôle dans la famille des protéines à MBD seraient très probablement MBD1 et MeCP2 et dans la famille Kaiso, Kaiso et ZBTB4, d'après leurs rôles et leurs redondances fonctionnelles potentielles et/ou avérées.

Comme il existe d'une part une corrélation directe entre les taux des transcrits *MBD1* et *MeCP2* et les taux protéiques et que d'autre part, les analyses des protéines par Western blot sont difficiles à réaliser à cause des anticorps employés, j'ai privilégié l'étude des transcrits *MeCP2* et *MBD1* au cours de la progression tumorale. Pour cela, j'ai réalisé une quantification de ces transcrits par PCR quantitative et normalisé les résultats par rapport au gène de ménage *PGK1*. Il apparaît que **le taux des transcrits MBD1 diminue fortement et de manière significative au cours de la progression tumorale**, avec une baisse de l'ordre de 5 fois entre HMEC et HMEC-*hTERT* et entre HMEC-*hTERT* et HMLER, soit de 25 fois entre les cellules normales et les cellules cancéreuses (**Figure 7A**). Par contre, **le taux de transcrits MeCP2 augmente au cours de la progression tumorale**, avec une hausse significative de l'ordre de 4 fois entre HMEC et HMEC-*hTERT* et de 1,1 fois donc de façon **non significative entre HMEC-*hTERT* et HMLER (Figure 7B)**. De plus, ces taux ne paraissent pas plus élevés que dans les cellules HeLa (résultat non représenté) dans lesquelles la protéine MeCP2 est peu abondante, ce qui révèle **que les taux protéiques de MeCP2 sont relativement faibles dans le modèle de progression tumorale**. Toutes ces données ne **semblent pas faire de MBD1 ou de MeCP2 de bons candidats** pour le remplacement de la protéine MBD2.

Concernant les deux membres de la famille Kaiso, une analyse de leur taux protéique a été réalisée par Western Blot, Ku80 servant de contrôle de dépôt. **Le taux de Kaiso semble constant entre HMEC-*hTERT* et HMLER alors que le taux de ZBTB4 augmente (Figure 7C)**. De plus, la comparaison avec les cellules HeLa dans lesquelles ces protéines sont abondantes, suggère que ces **protéines sont fortement exprimées dans ces lignées**. Aussi, **Kaiso et plus particulièrement ZBTB4 paraissent être de bons candidats pour le remplacement de la protéine MBD2 dans les cellules HMLER**.

Conclusions et perspectives

Au cours de cette étude, nous avons montré que les modifications de la méthylation de l'ADN (dans la zone comprise entre -700 et +200 pb après le site d'initiation de la transcription) sont limitées au cours de la progression tumorale, dans notre modèle syngénique de progression tumorale mammaire. En revanche, en cherchant parmi les différents paramètres étudiés, lequel pouvait être le plus discriminant entre les cellules immortalisées et les cellules tumorales, nous avons mis en évidence une forte redistribution de la protéine MBD2. Elle reste cependant un répresseur dépendant de la méthylation majeur au cours de la cancérogenèse, associé à une chromatine condensée.

La distribution différentielle de MBD2 entre les deux types cellulaires ne semble pas dépendre des modifications de la méthylation ou d'un changement de concentration protéique de MBD2. De plus, la comparaison des sites méthylés et silencieux entre les cellules immortalisées et tumorales suggère que MBD2 pourrait être remplacée, en fonction de la cellule considérée, par d'autres protéines liant l'ADN méthylé afin de conserver la répression transcriptionnelle tout au long de la progression tumorale. La redondance entre les protéines liant l'ADN méthylé reste à élucider mais dans ce modèle, Kaiso et ZBTB4 semblent être de bons candidats pour le remplacement de MBD2 entre HMEC-*hTERT* et HMLER. Aussi, des immunoprécipitations de chromatine avec des anticorps dirigés contre les protéines Kaiso et ZBTB4 viennent d'être effectuées et sont en cours d'analyse.

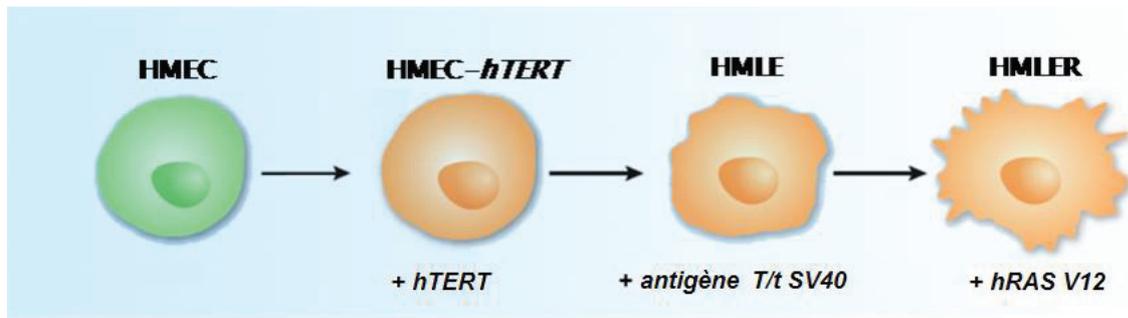


Figure 1: Modèle cellulaire syngénique de progression tumorale mammaire

Ce modèle utilise des cellules épithéliales mammaires humaines saines (HMEC) transfectées par infection virale (+) avec la transcriptase inverse de la télomérase humaine *hTERT* pour leur immortalisation (cellules HMEC-*hTERT*), les antigènes t et T de SV40 puis avec l'oncogène *Ras* humain version 12 (*hRAS-V12*) pour leur transformation (cellules HMLER)

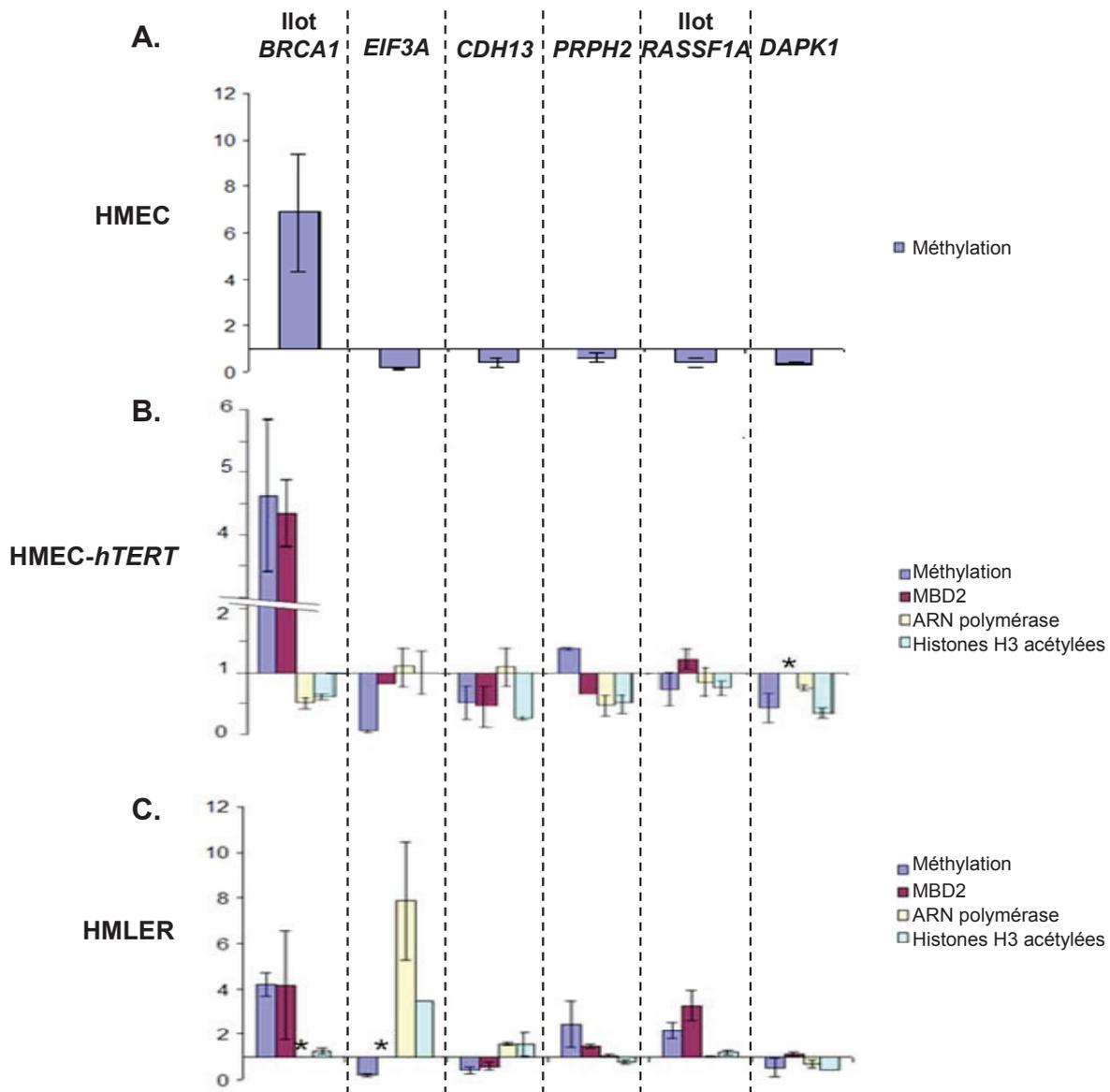


Figure 2 : Efficacités et spécificités des expériences de MeDP et ChIP

6 régions promotrices (inclues dans des îlots pour *BRCA1* et *RASSF1A* ou des régions moins riches en CpG pour *CDH13*, *PRPH2*, *DAPK1* et *EIF3A*) ont été analysées par PCR quantitatives pour les lignées cellulaires HMEC (A), HMEC-*hTERT* (B) et HMLER (C). Pour chaque expérience et chaque région, la fraction précipitée d'ADN ou de chromatine a été comparée à la fraction initiale. Chaque barre d'histogramme représente donc la moyenne \pm écart moyen de 3 expériences indépendantes (sauf pour les régions marquées * : données non disponibles). La région « îlot *BRCA1* » sert de contrôle positif et *CDH13* de contrôle négatif pour les MeDP et les ChIP anti-MBD2. Les autres régions représentent un panel concernant les profils de méthylation de l'ADN (barres bleues), de présence ou d'absence de la protéine MBD2 (en violet), de l'ARN polymérase II (en jaune), d'histones H3 acétylés (en turquoise). Ces régions ont servi à établir les seuils pour chaque paramètre dans chaque lignée cellulaire.

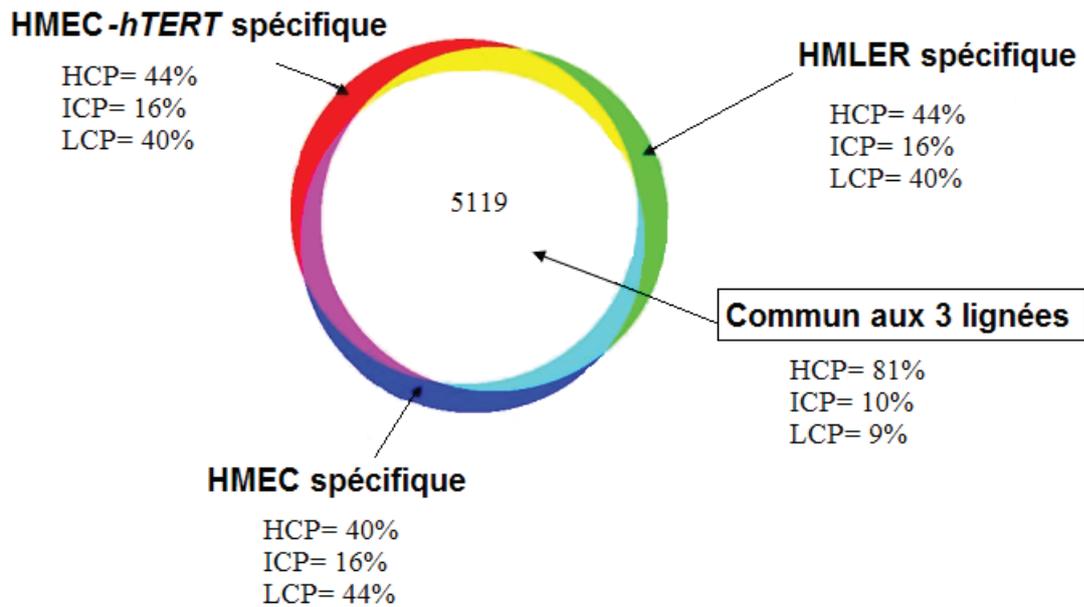


Figure 3: Diagramme de Venn représentant la distribution des sites méthylés au cours de la progression tumorale et classification des promoteurs en fonction de leur C+G et fréquence en CpG.

Les promoteurs présentant au moins un site méthylé dans la région comprise entre -700 et +200 pb du site d'initiation de la transcription ont été croisés entre les cellules HMEC, HMEC-*hTERT* et HMLER. Des sites sont exclusivement méthylés pour chaque lignée et 5119 sites paraissent communs aux 3 lignées. Pour chacun de ces cas, la répartition des promoteurs en fonction de leur richesse en CpG (HCP), pauvreté en CpG (LCP) et de densité intermédiaire (ICP) est indiquée en pourcentages.

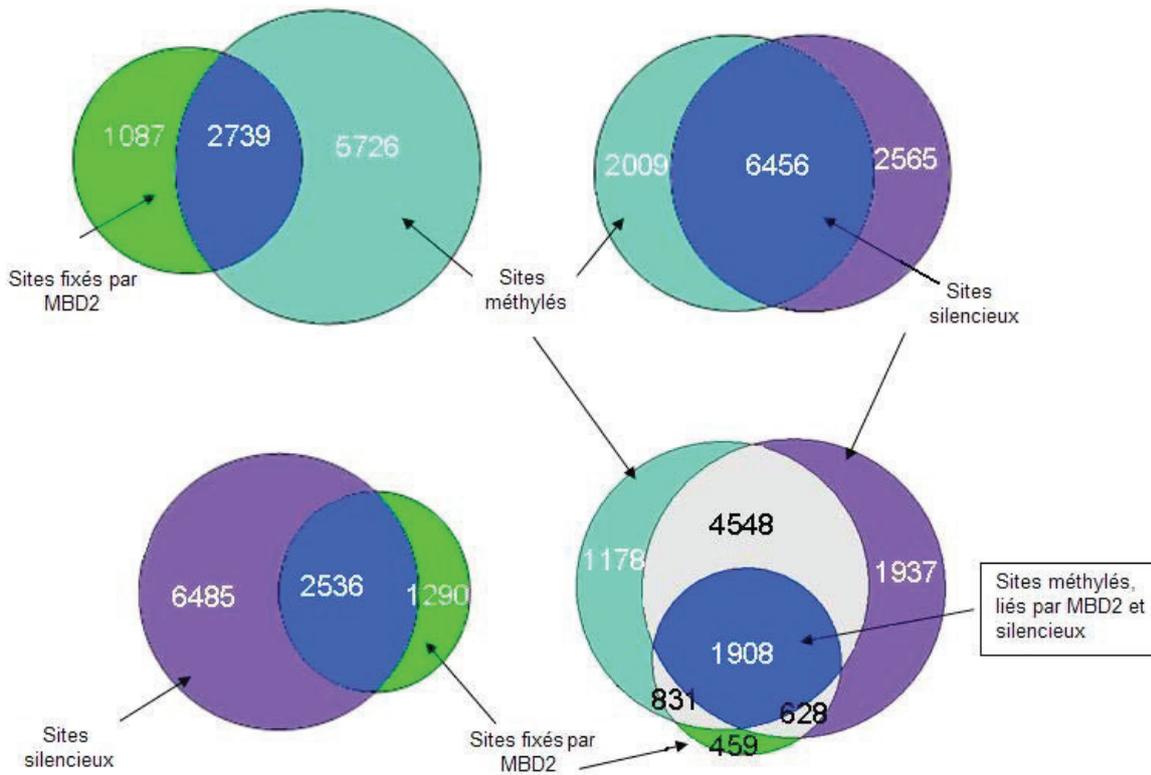


Figure 4: Diagrammes de Venn représentant la répartition des sites méthylés, fixés par la protéine MBD2 et non liés par l'ARN polymérase II (gènes silencieux) dans les cellules HMEC-*hTERT*.

Les sites liés par MBD2, représentant ~72% des sites méthylés (A), sont associés à 76% (B) à une répression transcriptionnelle. Une forte corrélation est établie entre la méthylation de l'ADN (76%) et la répression transcriptionnelle (72%) (C). Les nombres indiqués correspondent pour la partie bleue aux sites communs et pour le reste aux sites non superposés. Les tests χ^2 réalisés pour les croisements A, B et C (respectivement $p= 1,98 \times 10^{-8}$; $p= 4,09 \times 10^{-42}$; $p= 2 \times 10^{-25}$) montrent une corrélation significative entre les paramètres. La comparaison de toutes ces informations montre que 42% des sites (1908) MBD2 sont méthylés et correspondent à des gènes silencieux (D).

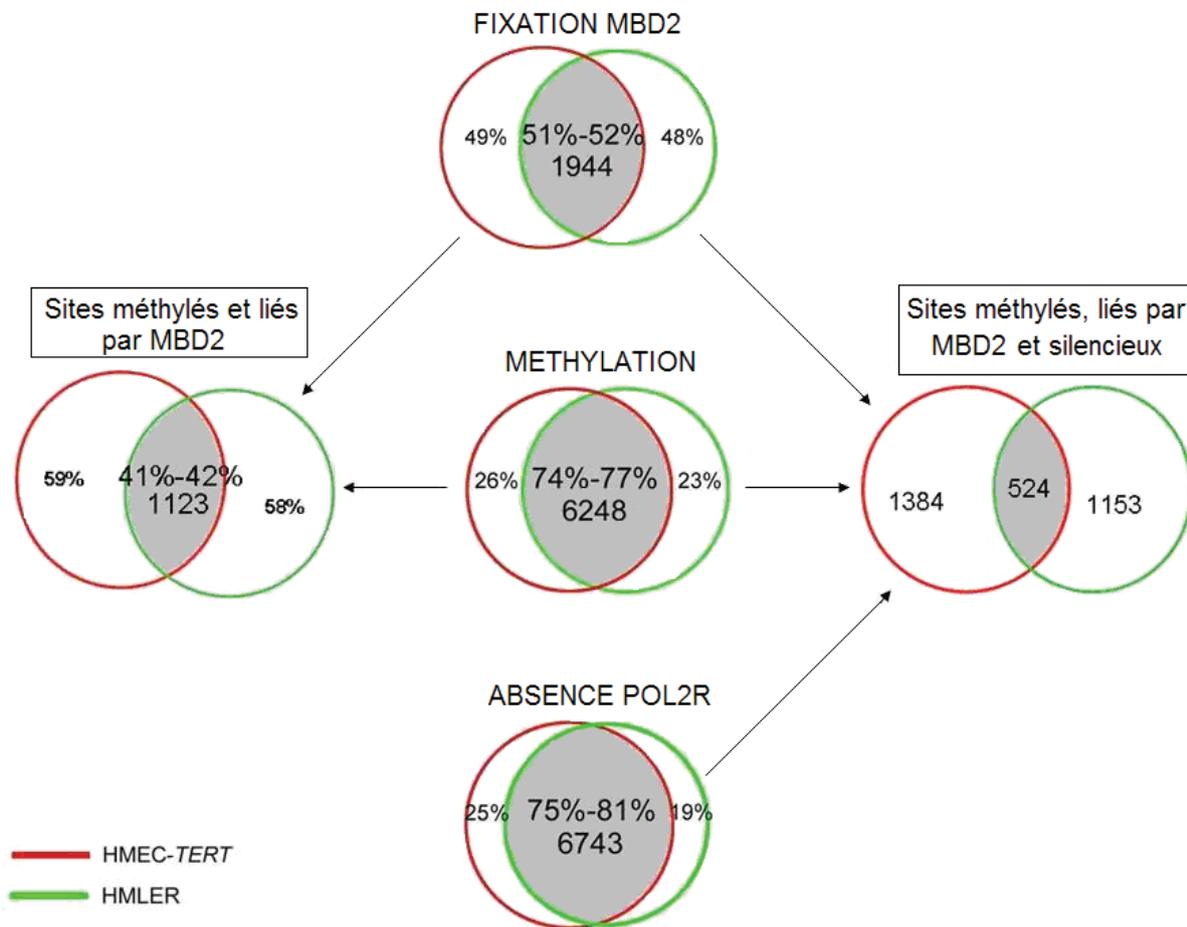


Figure 5: Diagrammes de Venn représentant entre HMEC-*hTERT* et HMLER la répartition des promoteurs selon les paramètres de fixation de la protéine MBD2, de méthylation de l'ADN ou d'absence de la polymérase II.

Au centre se trouvent les comparaisons de chaque paramètre entre les cellules HMEC-*hTERT* (en rouge) et les cellules HMLER (en vert). Les pourcentages indiquent pour chaque lignée le taux de superposition (dans la partie grise) ou de spécificité (dans la partie extérieure). Le paramètre le plus discriminant est la fixation de la protéine MBD2 et la prise en compte du paramètre méthylation (à gauche) ne change pas sensiblement les pourcentages. A droite, il apparaît que les lignées cellulaires HMEC-*hTERT* et HMLER partagent peu (524) de gènes silencieux avec des promoteurs méthylés et fixés par la protéine MBD2.

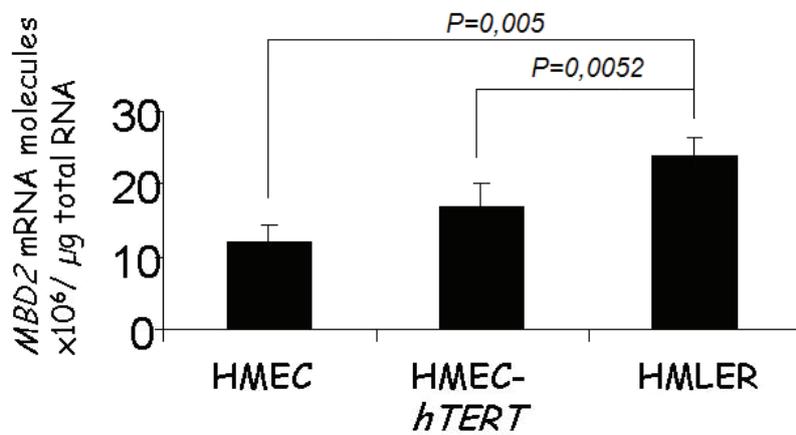


Figure 6: Quantification des transcrits *MBD2* dans le modèle syngénique de cancer mammaire.

Le dosage des transcrits *MBD2* s'est effectué par RT-PCR compétitive comme décrit précédemment (ref), à partir de 3 échantillons indépendants pour chaque lignée HMEC, HMEC-*hTERT* et HMLER. Chaque barre de l'histogramme correspond à la moyenne \pm écart moyen du nombre de transcrits *MBD2*, exprimé en nombre de molécules d'ARNm *MBD2* par microgramme d'ARN total. Les valeurs P indiquées sont issues des tests t effectués entre HMLER et les autres lignées du modèle.

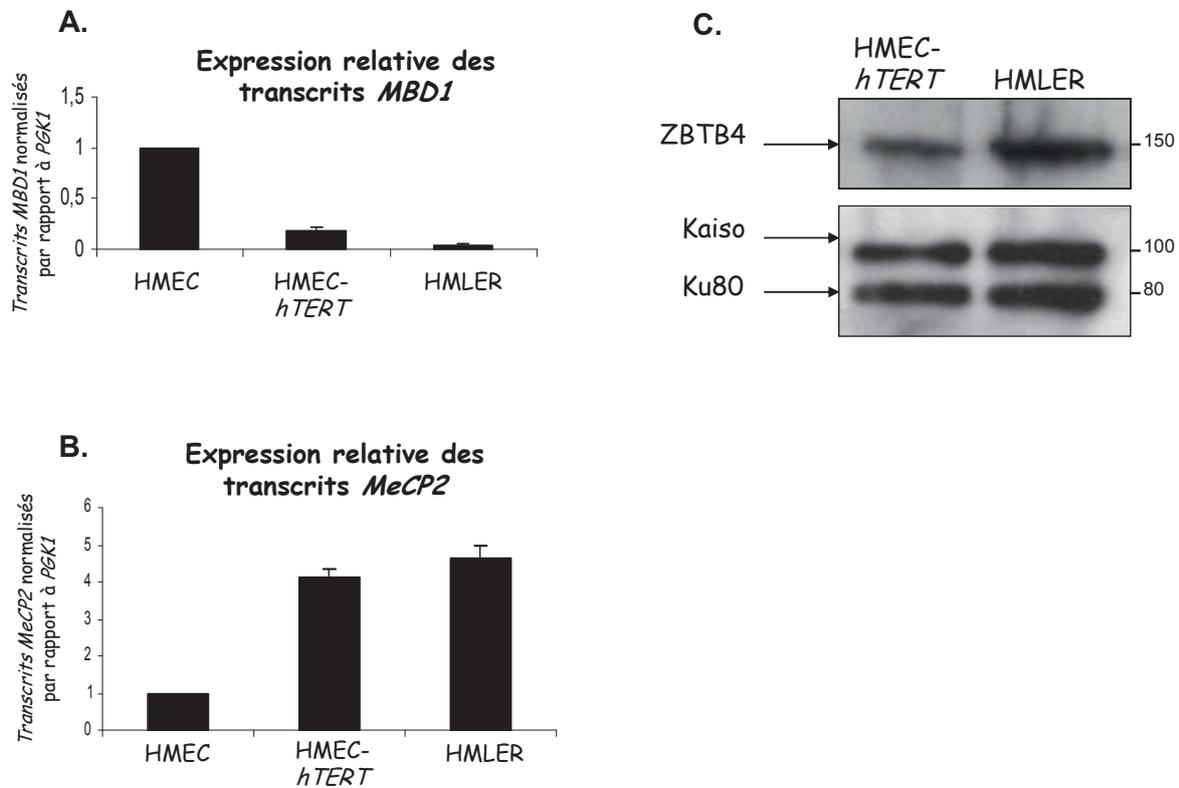


Figure 7: Quantifications des transcrits *MBD1* et *MeCP2* et des protéines Kaiso et ZBTB4.

Les dosages des transcrits *MBD1* (A) et *MeCP2* (B) ont été effectués par PCR quantitative, à partir de 3 échantillons indépendants pour chaque lignée, puis normalisés par rapport au gène de ménage *PGK1*. Chaque barre de l'histogramme correspond à la moyenne \pm écart moyen du nombre de transcrits *MBD1* ou *MeCP2*, rapportés aux ratios déterminés pour HMEC.

Quantification protéique par Western Blot des protéines ZBTB4 et Kaiso dans les lignées HMEC-*hTERT* et HMLER, Ku80 servant de témoin de dépôt (C).

		séquence dans le sens 5' --> 3'	position par rapport au TSS	Température d'hybridation
ilot <i>BRCA1</i>	sens	AAG GGC TCC TCC AGC ACG GC	de -1364 à -1218	66°C
	antisens	TTC TGA GGG ACC GAG TGG GC		
<i>CDH13</i>	sens	GAA GGC TGC ATG GAA ATG AT	de -828 à -606	61,5°C
	antisens	GGA TCT GGA TCG GAA GGA A		
<i>PRPH2</i>	sens	CCA CCA CTA CTG CAT TCA CG	de +51 à +231	62°C
	antisens	CAA GTG TAG TCC GGG TTG CT		
ilot <i>RASSF1A</i>	sens	TGT CCT CGG AGT CCA TCT TC	de +677 à +907	55°C
	antisens	GTG GTC GCG TGC ATA TTT TA		
<i>DAPK1</i>	sens	CAC CAG AAT GCA GGT TTC C	de -424 à -260	60°C
	antisens	CTA AGC TGC CCT CTG ACT GG		
<i>EIF3A</i>	sens	TGA CGA CAC TCC TTT CCG TG	de -638 à -332	66°C
	antisens	CCT GGG AAT CGC CTG TCT TG		

Tableau 1: Amorces utilisées pour l'étude des régions promotrices

Pour chaque amorce sont indiqués le nom, le sens, la position par rapport au site d'initiation de la transcription (TSS) et la température d'hybridation.

	lignées	seuils	nombre de gènes
Méthylation de l'ADN	HMEC	0,1	8416
	HMEC-<i>hTERT</i>	0,1	8465
	HMLER	0,1	8104
Liaison de MBD2	HMEC-<i>hTERT</i>	0,5	3826
	HMLER	0,5	3736
Présence d'histones acétylées	HMEC-<i>hTERT</i>	0,4	2679
	HMLER	0,3	2879
Absence d'ARN polymérase II	HMEC-<i>hTERT</i>	<-0,1	9021
	HMLER	<-0,1	8353

Tableau 2 : Nombre de gènes associés à chaque paramètre en fonction du type cellulaire et du seuil fixé

	catégorie	valeur p (GOTM)	nombre de gènes
processus biologique	régulation de la transcription (GO:0045449)	0,0004	89/451
fonction moléculaire	liaison à l'ADN (GO:0003677)	0,0005	86/451
	liaison aux protéines (GO:0005515)	0,0024	215/451

Tableau 3 : Fonctions majeures des protéines codées par les gènes présentant des promoteurs méthylés, liés par la protéine MBD2 et associés à une répression transcriptionnelle dans les cellules HMEC-*hTERT* et HMLER

DÍSCUSSION

1. Mise au point d'un outil permettant la sélection spécifique de régions d'ADN méthylées quelque soit leur densité en doublets CpG

Le profil de méthylation de l'ADN d'une région d'intérêt peut être déterminé par modification au bisulfite de sodium suivi d'un séquençage. Nous avons ainsi mis en évidence par cette technique que les promoteurs des gènes *LIMS2*, *EGFL7*, *PRPH2*, *FAM65A*, *DNASE1* et *IL2RB* étaient méthylés dans les cellules MRC5.

Si cette technique permet d'identifier pour chaque CpG son profil de méthylation, elle est inadaptée pour réaliser une cartographie des profils de méthylation du génome et la comparer aux cibles spécifiques de la protéine MBD2. Aussi, nous avons besoin de développer un outil permettant de sélectionner spécifiquement les régions de l'ADN méthylées. Cette sélection devait se faire quelque soit la densité en CpG puisque si la protéine MBD2 reconnaît des régions riches en CpG méthylés (îlot *BRCA1* (Auriol, Billard et al. 2005)), elle est aussi capable de fixer des régions promotrices méthylées pauvres en CpG (*pS2* dans les cellules HeLa (Chatagnon, Ballestar et al.)).

Etant donné que l'immunoprécipitation de chromatine avec des anticorps dirigés contre la 5-méthyl-cytosine nécessite une dénaturation préalable de l'ADN afin de le rendre simple brin et reconnaissable par l'anticorps, que la précipitation favorise les zones riches en CpG méthylés et que l'anticorps peut présenter des variations de spécificité selon le lot utilisé, cette technique ne convenait pas pour notre étude. Nous nous sommes alors intéressés à l'utilisation de protéines recombinantes comprenant des domaines MBD afin de sélectionner l'ADN méthylé. Au moment où nous avons commencé ce projet, il existait les protéines recombinantes développées par Jørgensen et ses collaborateurs (Jørgensen, Adie et al. 2006), comprenant des domaines MBD de MBD1, celles constituées de domaines MBD de MeCP2 (Cross, Charlton et al. 1994) ou de MBD2 (Yegnasubramanian, Lin et al. 2006). Comme une spécificité de séquence avait été décrite pour MeCP2 et que le domaine MBD de MBD2 est inclus dans le domaine TRD, notre choix s'est porté sur les protéines chimériques comprenant le domaine MBD de MBD1, étiquetées par 6 histidines consécutives et par un groupement HA, et pouvant se lier à des séquences nucléotidiques contenant un ou plusieurs CpG méthylés. Nous avons privilégié la protéine constituée de 4 domaines (protéine 4xMBD) car elle se fixe aux séquences méthylées avec une affinité au moins 50 fois plus grande par rapport à la protéine contenant 1 seul domaine.

Nous avons réalisé une production de la protéine 4xMBD (et de son équivalent non fonctionnel muté dans tous les domaines MBD au niveau de l'acide aminé 22) à partir du vecteur plasmidique fourni par cette équipe et mis au point les conditions du milieu protéique afin que la protéine 4xMBD ne se dégrade pas, reste soluble et fonctionnelle c'est-à-dire liée spécifiquement à des séquences méthylées.

Puis, j'ai mis au point une technique de précipitation d'ADN méthylé. Cette méthode présente l'avantage de fixer l'ADN natif, de partir de petites quantités d'ADN (500ng) donc de cellules et de disposer d'un outil reproductible. Nous avons mis en évidence par PCR quantitative, en comparant les fractions précipitées aux fractions initiales, que la protéine 4xMBD précipite spécifiquement des régions méthylées, qu'elles soient pauvres ou riches en CpG (*pS2*/îlot *BRCAl*) avec un taux d'enrichissement similaire, montrant ainsi que cette technique est adaptée pour l'identification des régions méthylées, sans biais dû à la densité en CpG. Cette précipitation d'ADN méthylé (MeDP), couplée à l'hybridation sur puces promoteurs, permet ainsi de déterminer les profils globaux de méthylation de l'ADN dans tout modèle cellulaire. Elle paraît même être envisageable pour l'identification de régions méthylées dans des échantillons tumoraux, à condition de l'ajuster pour des quantités d'ADN encore plus faibles, hypothèse que nous testons actuellement dans l'équipe.

2. La protéine MBD2, un répresseur transcriptionnel dépendant de la méthylation de l'ADN

2.1 Caractéristiques des sites fixés par la protéine MBD2

2.1.1 La protéine MBD2 lie spécifiquement des régions méthylées

La technique d'immunoprécipitation de chromatine (ChIP) nous a permis de caractériser *in vivo* les cibles directement liées par la protéine MBD2. Ainsi, « par approche gène candidat » nous avons pu par exemple détecter la présence de cette protéine sur les promoteurs méthylés des gènes *BRCAl* dans toutes les lignées cellulaires étudiées, *hTERT* dans différentes lignées cellulaires cancéreuses, *pS2* dans les cellules cancéreuses HeLa, *RASSF1A* dans les cellules transformées (HMLER) de notre modèle de progression tumorale mammaire, *LIMS2*, *EGFL7*, *PRPH2*, *FAM65A*, *DNASE1* et *IL2RB* dans les cellules normales MRC5.

Ces liaisons apparaissent spécifiques. En effet, dans les cellules HeLa, nous avons vérifié que certaines zones connues pour être non méthylées et non fixées par la protéine MBD2 (comme l'exon1 de *BRCAl* (Auriol, Billard et al. 2005)) gardaient ce profil dans nos CHIP anti-MBD2. Pour les cellules MRC5, la zone choisie dans le promoteur *CDH13* ne contenait pas de CpG et effectivement aucun enrichissement n'a été constaté. Pour le modèle de progression tumorale mammaire, c'est la répétition des expériences qui a montré cette spécificité.

Dans les cellules MRC5 et notre modèle de progression tumorale, nous n'avons pas testé si cette fixation était exclusive. En revanche, dans les cellules HeLa, nous avons montré pour le gène *hTERT* que parmi les protéines liant l'ADN méthylé, ni MeCP2 ni MBD1 n'étaient retrouvées sur la région promotrice méthylée. Cette étude fait suite à celles déjà effectuées dans ces cellules pour le gène *pS2* (Chatagnon, Ballestar et al.) et le gène *BRCAl* (Auriol, Billard et al. 2005) où MBD2 paraît aussi être la seule de ces 3 protéines à MBD à lier leurs promoteurs méthylés. Pour ces gènes, nous sommes donc plutôt dans la situation « 1 MBD par gène ». La perte de MBD2 par transfections transitoires ou stables d'ARN interférants dirigés contre l'ARNm *MBD2* ne suffit pas à induire le remplacement de la protéine MBD2 par MeCP2 ou MBD1. Aussi, pour ces gènes, le recrutement de MBD2 semble spécifique et ne pas dépendre de la concentration de ces protéines.

2.1.2 La protéine MBD2 se concentre dans des régions comprises entre -700 et +200 pb autour du site d'initiation de la transcription

Nous avons mis au point une technique de « CHIP-on-chip » permettant de suivre la répartition des protéines MBD2 le long des promoteurs, sur des régions comprises entre -7,45kb et +2,45kb autour du site d'initiation de la transcription. Un pic de localisation apparaît dans les régions situées entre -700 et +200pb autour du site d'initiation de la transcription (TSS). Ces résultats suggèrent donc que la protéine MBD2 joue un rôle essentiel dans la répression de l'initiation de la transcription. Par contre, notre approche ne permet pas de déterminer si la protéine MBD2 intervient dans l'élongation.

Nous avons utilisé cet intervalle (-700pb ; +200pb) pour comparer les données issues de « CHIP antiMBD2 -on-chip » et de « MeDP-on-chip », ceci afin de superposer les profils globaux de fixation de la protéine MBD2 à ceux de la méthylation de l'ADN d'une part dans les cellules HeLa et d'autre part dans le modèle de progression tumorale. Nous avons confirmé dans chaque lignée les liaisons observées par approche « gène candidat ». Nous avons aussi pu déterminer qu'une forte proportion de promoteurs liés par MBD2 était méthylée, c'est-à-dire 74% dans les cellules HeLa, 71% et 72% respectivement dans les cellules immortalisées et transformées de notre modèle de progression tumorale mammaire. Aussi, la méthylation de l'ADN semble jouer un rôle majeur dans la fixation de la protéine MBD2 dans ces cellules différentes par rapport au type cellulaire (col de l'utérus, sein), à la morphologie (épithéliales pour HeLa et HMEC-*hTERT*, mésenchymateuses pour HMLER), et au stade de progression tumorale (précoce pour HMEC-*hTERT*, tardif c'est-à-dire cancéreux pour HMLER et HeLa). Ces pourcentages signifient aussi qu'il existe des sites fixés par la protéine MBD2 qui ne sont pas méthylés. Ceci peut être être le reflet d'un recrutement de la protéine MBD2 par d'autres protéines sur ces sites non méthylés, bien que l'on ne puisse exclure des défauts techniques avec certains sites méthylés non identifiés par « MeDP-on-chip » et/ou de quelques faux positifs apparus en « CHIP antiMBD2-on-chip ».

2.1.3 La liaison de MBD2 ne dépend pas de la densité en CpG

Nous avons défini dans les cellules HeLa, par une approche systémique, que 6108 sites sont méthylés et fixés par la protéine MBD2. En utilisant les critères d'un îlot de CpG définis par Takai et ses collaborateurs (Takai and Jones 2002) à savoir une région de > 500pb avec un taux C+G $\geq 55\%$ et un ratio $CpG_{o/t} > 0,65$, nous avons pu déterminer par le logiciel CpGProd qu'environ la moitié de ces 6108 régions présentait un îlot de CpG dans la zone comprise entre -700 et +200pb autour du TSS. Ceci signifie donc que la densité en CpG n'est pas un facteur discriminant pour la fixation de la protéine MBD2 pour ces cellules.

Dans le modèle de progression tumorale, nous avons réalisé une analyse plus fine des promoteurs selon les critères définis par Weber et ses collaborateurs (Weber, Hellmann et al. 2007). Un promoteur est riche en CpG (HCP) si la région de > 500pb répond à une fréquence > 55% et un ratio > 0,75, pauvre en CpG (LCP) si le ratio est < 0,48 ou classé intermédiaire (ICP) si ce n'est ni un HCP ni un LCP.

Nous avons ainsi établi pour les cellules immortalisées et transformées un classement des promoteurs d'une part méthylés et d'autre part liés par la protéine MBD2. Il apparaît que pour les deux lignées cellulaires, la distribution des promoteurs fixés par la protéine MBD2 suit celle des promoteurs méthylés (de l'ordre de 72% pour les HCP, 11% pour les ICP et 17% pour les LCP), montrant une nouvelle fois qu'il n'y a pas de sélection préférentielle de régions méthylées en fonction de leur densité en CpG pour la liaison de la protéine MBD2.

2.1.4 La protéine MBD2 ne se lie à aucune séquence consensus

Nous avons cherché à identifier par le logiciel CisGenome un éventuel motif de séquence consensus pour la liaison de la protéine MBD2. Nous n'avons révélé qu'un motif CG voire peut être CCG. Aussi aucune spécificité de séquence n'a pu être rapportée pour la protéine MBD2, contrairement à d'autres protéines à MBD. En effet, MeCP2 lierait préférentiellement des sites CCGG méthylés présents dans des séquences riches en A/T (Klose, Sarraf et al. 2005) alors que la fixation de MBD1 sur un CpG méthylé paraît favorisée en présence d'autres CpG méthylés successifs ou lorsqu'il est inclus dans un motif (TouTG)CGCA (Clouaire, de Las Heras et al.).

2.2 La protéine MBD2 joue un rôle majeur dans la répression transcriptionnelle

Au cours des études « gènes candidats », nous avons pu montrer par RT-PCR quantitative que la protéine MBD2 participe à leur répression transcriptionnelle. Si ce rôle avait déjà été identifié (Chatagnon, Ballestar et al.; Ballestar, Paz et al. 2003; Auriol, Billard et al. 2005), d'autres études laissaient entendre que cette protéine pouvait aussi intervenir comme activateur transcriptionnel (Fujita, Fujii et al. 2003; Lembo, Pero et al. 2003; Angrisano, Lembo et al. 2006). Aussi, pour caractériser le rôle de cette protéine, nous avons réalisé un croisement des profils de méthylation de l'ADN, de liaison de la protéine MBD2 avec les listes de gènes silencieux issues de « ChIP-on-chip » montrant la fixation de l'ARN polymérase II. Il apparaît que les sites fixés par la protéine MBD2 sont associés à une inhibition transcriptionnelle à 65%, 66%, 59% respectivement dans les cellules HeLa,

HMEC-*hTERT* et HMLER et que 51%, 42% et 45% des promoteurs liés par la protéine MBD2 sont méthylés et associés à une répression transcriptionnelle. Tous ces résultats montrent donc que la protéine MBD2 est un vrai répresseur très majoritairement dépendant de la méthylation de l'ADN. Il apparaît également que les gènes méthylés et silencieux sont fixés par la protéine MBD2 à 41%, 30% et 29% respectivement dans les cellules HeLa, HMEC-*hTERT* et HMLER. Ceci suggère donc que la protéine MBD2 est un répresseur majeur, peut être le plus important parmi toutes les protéines liant l'ADN méthylé qui se partagent, avec tous les mécanismes impliquant directement la méthylation de l'ADN, le reste de l'inhibition transcriptionnelle dépendante de la méthylation de l'ADN. Ces résultats avaient déjà été suggérés pour les cellules HeLa, suite à l'étude de l'impact de l'absence des protéines MBD2, MeCP2 et MBD1 sur l'expression génique, au cours de laquelle MBD2 apparaissait réprimer le plus grand nombre de gènes (Lopez-Serra, Ballestar et al. 2008).

Enfin, dans les cellules cancéreuses HeLa, nous avons également mis en évidence que 12% des promoteurs liés par la protéine MBD2 sont associés à la présence de l'ARN polymérase II. Ceci montre également que l'éventuel rôle activateur de la protéine MBD2 serait très minoritaire.

2.3 La répression dépendante de la méthylation initiée par la protéine MBD2 est associée à une compaction chromatinienne

Nous avons vu au cours de l'introduction, au cours du chapitre 2, que la répression transcriptionnelle initiée par la protéine MBD2 était principalement dépendante du recrutement des désacétylases d'histones. Aussi, dans le modèle de progression tumorale mammaire, nous avons pu suivre la distribution des sites acétylés sur l'histone H3, une marque d'activation transcriptionnelle. Il apparaît que, dans les cellules HMEC-*hTERT* et HMLER, plus de $\frac{3}{4}$ des sites méthylés, réprimés suite à la fixation de la protéine MBD2 ne sont pas associés à l'acétylation des histones H3. Ces données laissent donc supposer en accord avec la littérature, qu'à l'échelle du génome la protéine MBD2 initie la répression transcriptionnelle dépendante de la méthylation en recrutant majoritairement des complexes de désacétylation des histones remodelant la chromatine afin de la condenser et de la rendre inaccessible à la machinerie transcriptionnelle. Pour le reste des sites associés à une

acétylation des histones, ceci pourrait signifier que ces sites deviennent silencieux au cours de la progression tumorale et que le recrutement et/ou l'action des HDAC n'ont pas encore eu le temps de se mettre en place. Ces régions promotrices pourraient aussi présenter une différence en nombre de CpG méthylés induisant une différence de compaction de la chromatine. En effet, les données de Curradi et ses collaborateurs montrent que la méthylation de peu de cytosines peut mener à l'inhibition transcriptionnelle mais qu'un certain nombre de CpG méthylés est nécessaire pour qu'une compaction de la chromatine ait lieu via notamment la désacétylation des histones (Curradi, Izzo et al. 2002).

3. La force de répression de la protéine MBD2 est modulable

3.1. La force de répression de MBD2 dépend de la distance au TSS

La majorité des sites méthylés, fixés par MBD2 et associés à la répression transcriptionnelle se superposent dans la région comprise entre -700 et +200 pb autour du site d'initiation de la transcription (TSS), suggérant que cette zone est importante pour l'inhibition de l'expression génique. C'est également dans cet intervalle que la corrélation entre méthylation de l'ADN et répression transcriptionnelle paraît la plus étroite (Weber, Hellmann et al. 2007). La force de répression de MBD2 serait donc la plus importante dans cet intervalle.

Cette distance d'action est assez semblable à celle de la protéine MeCP2. En effet, il apparaît que cette protéine peut intervenir au minimum à 500pb en amont du TSS (Nan, Campoy et al. 1997; Kaludov and Wolffe 2000). On peut donc supposer que la région comprise entre -700 et +200pb autour du TSS est la zone essentielle à l'initiation de la répression transcriptionnelle initiée par la liaison des protéines à MBD sur l'ADN méthylé.

Néanmoins, il existe aussi une petite proportion de sites liés par la protéine MBD2, méthylés et réprimés situés sur le reste de la région promotrice comprise entre -7,450 à +2,450 pb autour du TSS (limites de la puce « promoteurs »). Le nombre de gènes réprimés associés à des promoteurs méthylés liés par la protéine MBD2 diminue graduellement en fonction de la distance par rapport au TSS. Ces résultats suggèrent également que la répression transcriptionnelle initiée par MBD2 est liée à la distance. De plus, une distance importante ne signifie pas pour autant absence de répression. En effet, il a été démontré par

l'étude du locus *GATA-4* que MBD2 pouvait participer à une répression transcriptionnelle grâce à sa présence dans une structure en boucle (Tiwari, McGarvey et al. 2008) donc sur des sites éloignés du TSS.

3.2 La force de répression de MBD2 dépend de la densité en CpG

La technique d'interférence à l'ARN dirigée contre l'ARNm *MBD2* couplée à l'analyse globale de l'expression génique en utilisant des puces transcriptomiques nous a permis d'évaluer l'impact de l'absence de MBD2 sur la régulation transcriptionnelle. Dans les cellules normales fibroblastiques embryonnaires de poumon MRC5, nous avons ainsi démontré que la baisse de concentration de la protéine MBD2 induit majoritairement la ré-expression de gènes présentant des promoteurs méthylés de concentration notable en CpG (promoteurs ICP) alors qu'elle n'entraîne qu'une stimulation de l'expression de gènes présentant des promoteurs riches en CpG méthylés (promoteurs HCP). La densité en CpG des promoteurs méthylés semble donc être une composante importante dans la force de répression de la protéine MBD2. Ces résultats supposent aussi l'existence d'une ou d'autres « couche(s) » épigénétiques qui se superposeraient au mécanisme de répression initiée par la protéine MBD2 afin d'empêcher l'activation transcriptionnelle de gènes avec des HCP. On pourrait alors penser que la chromatine est plus condensée au niveau des promoteurs HCP que des ICP. Leurs différences concernant la réexpression génique pourrait être expliquée par un temps d'inhibition de la protéine MBD2 insuffisant pour induire un changement chromatinien assez avancé pour une activation transcriptionnelle, contrairement aux promoteurs ICP qui présentaient un état déjà plus décondensé donc plus favorable. Les promoteurs HCP pourraient aussi présenter des modifications d'histones plus nombreuses ou plus diverses (Wang, Zang et al. 2008), ce qui rendrait la condensation plus complexe. L'absence de la protéine MBD2 ne pourrait alors provoquer qu'une décompaction limitée de la chromatine. Cette structure chromatinienne serait d'autant plus difficile à destabiliser si des protéines renforçant cette conformation étaient mises en jeu. Elles incluent les variants d'histones, les protéines HP1 liant les nucléosomes et les protéines polycomb (PcG). Ainsi, L3MBTL1 pourrait s'associer à HP1 γ , aux sites monométhylés H4K20 et mono ou di-méthylés H1bK26 afin de former un pont entre des nucléosomes adjacents, participant à la répression transcriptionnelle du gène *c-myc* dans les cellules HeLa (Trojer, Li et al. 2007). La protéine Ring1B remplirait le même type de fonction lorsqu'elle appartiendrait à un complexe

répresseur polycombe 1 like (PRC1-like), car sa présence est associée à une compaction de la chromatine induisant une répression transcriptionnelle de *Hox1*, sans que Ring1B ubiquitinye les queues d'histones (Eskeland, Leeb et al.).

D'autre part, de très grandes régions (~2Mb) très méthylées et sujettes aux modifications des queues d'histones sont également retrouvées, ce qui supposerait qu'un gène inclus dans ces zones soit plus difficilement accessible à la machinerie transcriptionnelle (Coolen, Stirzaker et al.). Toutes ces données de la littérature pourraient donc expliquer les différences observées entre les gènes à promoteurs ICP et HCP suite à la perte de la protéine MBD2.

4. La protéine MBD2 est redistribuée au cours de la progression tumorale

Nous avons analysé dans un modèle cellulaire syngénique humain de cancer du sein, l'évolution, au cours de la progression tumorale, de la méthylation de l'ADN, de la distribution des protéines MBD2, répresseurs transcriptionnel dépendant de la méthylation, de l'acétylation des histones ainsi que l'activité transcriptionnelle d'environ 12 000 promoteurs.

L'analyse des régions méthylées dans la zone du TSS (-700 ; +200) montre des modifications de la méthylation mais qui sont relativement limitées. En effet, en comparant le nombre de sites méthylés entre les cellules immortalisées HMEC-*hTERT* et les cellules transformées HMLER, il apparaît que le taux global de méthylation de l'ADN diminue de 4%. D'autre part, l'analyse de chaque promoteur a révélé que ~75% d'entre eux conservent leur méthylation au cours de la progression tumorale (par comparaison des cellules HMEC, HMEC-*hTERT* et HLMER).

Dans les cancers, l'hypométhylation est évaluée entre 10% et 30%. Si nos résultats ne sont pas directement comparables étant donné qu'une seule fraction du génome est étudiée (les régions proches des TSS), on peut quand même dire que l'ordre de grandeur reste similaire. Le taux que nous observons ne représente pas une valeur absolue, il répond plutôt à une combinaison d'hypo et d'hyperméthylations étant donné que nous avons ciblé diverses catégories de promoteurs. Pour mesurer plus précisément le taux global d'hypométhylation, il faudrait par exemple réaliser une chromatographie en phase liquide à haute performance pour analyser la composition en bases.

Nous avons alors recherché parmi les autres paramètres étudiés, lequel pouvait être le plus discriminant entre les cellules immortalisées et les cellules tumorales, dans le modèle étudié. Cette analyse a mis en évidence un évènement qui n'était pas prévisible *a priori*. En effet, ~75% des gènes restent silencieux au cours de la progression tumorale alors que seulement ~50% des promoteurs conservent une liaison de la protéine MBD2 lors du passage cellules immortalisées à cellules tumorales.

La distribution différentielle de MBD2 entre les deux types cellulaires ne paraît pas être la conséquence d'un changement de concentration de la protéine MBD2 au cours de la progression tumorale. En effet, un dosage des transcrits par RT-PCR compétitive a montré que le taux d'ARNm *MBD2*, corrélé au taux protéique, n'est pas significativement différent entre les cellules HMEC-*hTERT* et HMLER. Elle ne semble pas non plus dépendre des modifications de la méthylation puisque la prise en compte du paramètre « méthylation » change peu les pourcentages obtenus pour le paramètre « présence de MBD2 ». Aussi, contrairement à notre hypothèse de départ, les modifications de la méthylation de l'ADN ne sont pas la principale cause de la redistribution des protéines MBD2. Le recrutement de MBD2 paraît être guidé, sans doute par des facteurs nucléaires qu'il reste encore à identifier.

Une comparaison des profils de méthylation l'ADN, de liaison de l'ARN polymérase II, de présence d'acétylation sur les histones H3 avait été effectuée dans des régions promotrices riches en CpG entre la lignée cellulaire normale HMEC et des lignées cancéreuses mammaires qui ne dérivait pas de la lignée normale (Takeshima, Yamashita et al. 2009). Ils avaient alors déterminé que l'hyperméthylation touchait préférentiellement les gènes présentant initialement une faible transcription alors que les promoteurs initialement fixés par d'ARN polymérase II étaient plutôt exemptés. Nous avons pour notre part privilégié un modèle homogène de progression tumorale à partir de cellules HMEC, analysé des promoteurs quel que soit leur densité en CpG et complété ces analyses en suivant l'évolution de la distribution d'une protéine liant l'ADN méthylé au cours de la cancérogenèse.

5. La protéine MBD2, une piste séduisante pour un traitement anti-cancéreux

5.1 La répression de la protéine MBD2 est réversible

Nous avons établi par nos approches directes de ChIP anti-MBD2, couplées à l'étude de l'expression génique par RT-PCR quantitative et des profils de méthylation de l'ADN par modification au bisulfite de sodium que la protéine MBD2 exerce une répression transcriptionnelle directe, spécifique et dépendante de la méthylation de l'ADN. Cette inhibition transcriptionnelle paraît être levée, partiellement ou totalement, suite à la perte de la protéine MBD2 par interférence à l'ARNm *MBD2* dans des cellules cancéreuses HeLa ou normales MCR5. Ces études viennent renforcer les résultats déjà obtenus par d'autres approches gènes candidats (Chatagnon, Ballestar et al.; Auriol, Billard et al. 2005).

Le caractère réversible des marques épigénétiques est déjà exploité dans des thérapies déméthylantes, via l'inhibition des DNMT par exemple la 5-aza-2' desoxycytidine (décitabine), visant à réactiver les gènes suppresseurs de tumeurs. Cependant, l'inhibition globale et non ciblée de la méthylation de l'ADN serait un facteur d'instabilité chromosomique pour induire la réexpression génique (Eden, Gaudet et al. 2003). L'inactivation ciblée de la protéine MBD2 pourrait donc conduire à une levée plus sélective de la répression transcriptionnelle.

5.2 La perte de la protéine MBD2 aurait un impact limité sur la régulation transcriptionnelle

5.2.1 Un faible nombre de gènes touchés par l'absence de MBD2

Une alternative aux traitements déméthylants serait d'agir au niveau des protéines « interprètes » du signal de méthylation de l'ADN.

Nous avons montré dans nos études que la protéine MBD2 présente des spécificités de cibles. Nous avons également établi, par interférence à l'ARNm *MBD2* et analyse par puces transcriptomiques, qu'à l'échelle globale du génome l'inhibition de cette protéine ne semble toucher que quelques groupes de gènes, de l'ordre de 2% du génome dans les cellules MRC5. Ces résultats semblent directement à imputer à l'absence de la protéine MBD2 puisque les taux des transcrits *MBD1* et *MeCP2*, qui sont corrélés aux taux protéiques, restent similaires entre les cellules témoins et les cellules traitées. Il n'y a donc pas d'altération globale du

génomique. Ces données sont en accord avec celles obtenues chez la souris où il apparaît que la perte de *Mbd2*, que ce soit par invalidation (Hendrich, Guy et al. 2001) ou injection d'ARN antisens (Campbell, Bovenzi et al. 2004), n'a pas d'effet délétère et semble provoquer une altération phénotypique très limitée. Ces données pourraient constituer un paradoxe avec celles issues de « ChIP-on-chip » et « MeDP-on-chip » avec lesquelles nous avons établi que la méthylation de l'ADN inhibe la transcription en grande partie via MBD2. En fait, tous ces résultats suggèrent une redondance fonctionnelle, concomitante à des spécificités de cibles, entre la protéine MBD2 et d'autres protéines liant l'ADN méthylé afin de garantir la répression de certains gènes.

Ce type d'observations semble en adéquation avec les résultats obtenus chez la souris concernant la protéine *Mecp2*, protéine à MBD la plus abondante dans le cerveau. La comparaison du transcriptome de cerveaux de souris témoins ou invalidées pour *Mecp2* ne montre que quelques différences d'expression faibles mais significatives pour un nombre restreint de gènes (Tudor, Akbarian et al. 2002) alors que cette protéine paraît instaurer la répression d'un nombre important de gènes (Skene, Illingworth et al.).

5.2.2 Redondance fonctionnelle entre MBD2 et d'autres protéines liant l'ADN méthylé ?

Nous avons montré par recoupement des données « ChIP-on-chip » et « MeDP-on-chip » que 30% des promoteurs méthylés et silencieux sont des cibles de MBD2 dans les cellules immortalisées dans notre modèle de progression tumorale et que ce taux passe à 41% dans les cellules transformées (et dans les cellules cancéreuses HeLa).

Par contre seulement ~30% des promoteurs méthylés et réprimés par MBD2, dans les cellules immortalisées par *hTERT*, sont aussi réprimés par MBD2 dans les cellules transformées et vice-versa. Ceci suggère d'une part que l'inhibition de ces gènes participe à l'établissement du phénotype mammaire et d'autre part que les 70% restants présentent une différence dans leur méthylation et/ou dans la fixation de la protéine MBD2 et/ou dans la présence de l'ARN polymérase II entre les deux lignées cellulaires. En fait, il apparaît que 559 gènes dans HMEC-*hTERT* restent méthylés et silencieux au cours de la progression tumorale bien qu'ils perdent la protéine MBD2 dans les cellules HMLER alors que pendant ce temps 600 gènes gagnent MBD2. Ces résultats laissent donc aussi supposer qu'il existe une

redondance entre la protéine MBD2 et d'autres protéines liant l'ADN méthylé. Il ne peut pas non plus être exclu que plusieurs protéines liant l'ADN méthylé soient simultanément présentes sur ces sites, comme cela a déjà été démontré pour certains sites (Bakker, Lin et al. 2002), et que par conséquent la perte ou le gain de MBD2 n'ait pas d'incidence sur la régulation transcriptionnelle de ces gènes.

L'impact de l'inhibition de la protéine MBD2 sur la régulation génique serait limité mais non négligeable vu que seuls certains sites laissés vacants par MBD2 pourraient accueillir d'autres protéines liant l'ADN méthylé (Klose, Sarraf et al. 2005). Cette protéine est donc une cible de plus en plus attractive, d'autant plus qu'il a été mis en évidence qu'elle joue un rôle dans le développement et la croissance tumorale (Sansom, Berger et al. 2003; Berger and Bird 2005).

Enfin, il pourrait être intéressant de déterminer, toujours par une approche « ChIP-on-chip », les profils de liaison d'autres protéines liant l'ADN méthylé et notamment celles présentant un rôle important dans la répression transcriptionnelle afin d'une part d'évaluer le taux de redondance possible entre toutes ces protéines et d'autre part d'identifier les gènes qui resteront réprimés et ceux qui sont susceptibles d'être réexprimés suite à la perte de MBD2. Ces études pourraient compléter celles déjà effectuées par Ballestar et ses collaborateurs. En effet, ils avaient comparé par « ChIP-on-chip » les sites liés par MeCP2, MBD1, MBD2 et MBD3 dans des cellules cancéreuses mammaires mais à partir de puces présentant un nombre limité d'ilots de CpG (Ballestar, Paz et al. 2003). Ils avaient évalué à ~7% les sites présentant les 4 protéines à MBD, et ~24%, 10%, 9% et 6% les sites liés uniquement par MBD2, MBD3, MeCP2 and MBD1, respectivement. Pour l'instant, ce type d'analyses n'a pas été rapporté pour les membres de la famille Kaiso. Seule une étude suggère une redondance fonctionnelle entre les protéines à MBD et cette famille (Martin Caballero, Hansen et al. 2009). D'autre part, suivre l'évolution de la distribution des protéines liant l'ADN méthylé dans notre modèle homogène de progression tumorale permettrait aussi de déterminer si cette redondance est dépendante du stade de cancérogenèse.

Conclusions et perspectives

En définitif, pour identifier les cibles spécifiques de MBD2 et caractériser les conséquences fonctionnelles de ces liaisons sur la régulation génique, nous avons développé différentes approches globales dans des modèles cellulaires normaux, cancéreux et de progression tumorale.

Dans des cellules cancéreuses HeLa, nous avons analysé les profils de méthylation de l'ADN, de liaisons de MBD2 et de l'ARN polymérase II par « ChIP-on-chip » et « MeDP-on-chip » avec des puces promoteurs. Nous avons ainsi pu montrer que la protéine MBD2 possède toutes les caractéristiques d'un répresseur dépendant de la méthylation. De plus, elle apparaît comme le répresseur majoritaire dans ce mécanisme.

Ces mêmes approches couplées à l'analyse de l'acétylation des histones H3 ont été réalisées dans un modèle cellulaire syngénique de progression tumorale mammaire humaine. Une forte proportion de gènes silencieux et méthylés semble liée par la protéine MBD2. Ces résultats, couplés à ceux obtenus pour les cellules HeLa, suggèrent donc que la méthylation de l'ADN réprime la transcription en grande partie via MBD2, et que ce mécanisme n'est pas spécifique d'un type cellulaire ou d'un stade de progression tumorale.

D'autre part, les comparaisons entre cellules immortalisées et transformées n'ont montré la modification de la méthylation de l'ADN que d'un nombre limité de promoteurs alors qu'une forte redistribution de MBD2 parmi ces sites méthylés est observée. Cette absence de corrélation indique de façon surprenante que l'altération des profils de fixation de la protéine MBD2 ne suit pas celle des profils de la méthylation de l'ADN. Ces résultats, ajoutés à ceux montrant que la concentration protéique de MBD2 n'est pas modifiée entre ces 2 lignées, laissent supposer que la fixation de la protéine MBD2 est guidée. Aucun motif de séquence consensus ne semble se dégager. Aussi, il reste à identifier les éventuels facteurs nucléaires qui influenceraient la distribution de la protéine MBD2.

La redistribution de la protéine MBD2 n'est pas associée à des altérations importantes de la répression transcriptionnelle. Puisque plus de 1000 gènes restent méthylés et silencieux au cours de la progression tumorale quelque soit leur statut vis-à-vis de la présence de la protéine MBD2, il semble qu'il existe une redondance entre les protéines liant l'ADN méthylé. Depuis la découverte des protéines à domaine MBD, d'autres familles de répresseurs ont été identifiées : les membres de la famille Kaiso qui possèdent un domaine en doigt de zinc, ainsi que des protéines possédant un domaine SRA (SET and Ring finger-associated).

Ces protéines possèdent toutefois des caractéristiques bien distinctes. Ainsi, elles semblent intervenir dans la répression de voies de signalisation distinctes et ne pas toutes recruter les mêmes corépresseurs (Sasai and Defossez 2009). Elles représentent donc des possibilités de régulation transcriptionnelle différentes.

Notre prochain objectif est donc d'identifier les éventuels « remplaçants » de MBD2 et d'analyser les conséquences moléculaires (modifications des histones, répression stricte ou modulée...) et physiologiques (apoptose, sensibilité aux inhibiteurs de certaines voies de signalisation...) au cours de la progression tumorale. Cette étude est déjà initiée. Il apparaît que parmi les protéines candidates, les protéines à MBD, au moins MeCP2 et MBD1, sont, *a priori*, à exclure, car l'expression de MBD1 baisse progressivement entre HMEC, HMEC-*hTERT* et HMLER et l'expression de MeCP2 est très faible. Les protéines Kaiso et ZBTB4, membres de la famille Kaiso, semblent par contre intéressantes car elles sont bien exprimées dans les cellules HMEC-*hTERT* et HMLER. Des CHIP anti-Kaiso et anti-ZBTB4 ont été réalisés et leur analyse se fera soit par approche « gène candidat » soit par approche globale en utilisant des puces promoteurs (ChIP-on-chip) ou un séquençage à haut débit (ChIP-seq). En fonction des résultats obtenus, nous rechercherons, par des expériences d'interférence à l'ARN ciblant une ou des protéines liant l'ADN méthylé, à mettre évidence les conséquences potentielles sur les phénotypes cellulaires de la perte d'expression (combinée ou non) de ces protéines.

Nous pourrions également transposer ce type d'étude sur des biopsies de patients, réalisées à différents stades de progression tumorale mammaire afin de confirmer ou infirmer ces résultats.

D'autre part, nous avons étudié dans des cellules normales MRC5 l'impact sur l'expression génique de l'inhibition de MBD2 par interférence à l'ARN, en utilisant des puces transcriptomiques. La perte de MBD2 n'induit pas de surexpression génique globale, ce qui suggère aussi une redondance entre des protéines liant l'ADN méthylé. Parmi les gènes présentant une altération génique, $\frac{1}{4}$ semblent ré-exprimés et $\frac{3}{4}$ sur-exprimés. Nous pourrions donc compléter cette analyse par des « ChIP anti-MBD2 -on-chip » dans les cellules MCR5 pour 1) vérifier que les cibles réprimées possèdent très majoritairement des promoteurs ICP et 2) identifier si les sur-expressions sont directement ou indirectement dues à l'absence de la protéine MBD2. Si ces changements d'expression sont le reflet d'une perte directe de la protéine MBD2 sur les régions promotrices, cela supposerait que la protéine MBD2 pourrait aussi agir comme un modulateur ou que la répression observée pour *hTERT* via MBD2 est

loin d'être un cas isolé. Pour vérifier cette dernière hypothèse, selon le nombre de gènes concernés, il faudrait réaliser l'étude de la méthylation de l'ADN de chaque CpG présents dans leurs régions promotrices par approche « gène candidat » avec des modifications au bisulfite de sodium suivies d'un séquençage ou par approche globale par séquençage à haut débit, afin de savoir si ces promoteurs sont complètement méthylés, présentent une mosaïque de CpG méthylés et non méthylés ou encore s'ils incluent une zone non méthylée entourée de zones méthylées, comme c'est le cas pour *hTERT*.

Au cours de toutes ces études, nous avons également pu caractériser la force de répression de la protéine MBD2. La densité en CpG des promoteurs méthylés et la distance de fixation de la protéine MBD2 par rapport au site d'initiation de la transcription semblent être des composantes importantes. De plus, nos données suggèrent que la protéine MBD2 empêche l'initiation de la transcription. Il serait intéressant d'élargir nos études à des régions inter- ou intragéniques afin d'établir si la protéine MBD2 peut aussi agir sur l'élongation. Cela pourrait peut être élucidé, au moins en partie, comment la méthylation intragénique, qui semble représenter une part importante des sites méthylés (Weber, Davies et al. 2005), est impliquée dans la répression transcriptionnelle, comme le suggère un certain nombre de travaux. En effet, la protéine MeCP2 semble déjà une piste à écarter puisqu'à la suite de sa fixation sur des sites intragéniques, elle interviendrait dans l'activation transcriptionnelle (Yasui, Peddada et al. 2007).

Enfin, toutes ces études renforcent l'idée que la protéine MBD2 constitue une cible intéressante pour une thérapie anti-cancéreuse. Son inhibition permettrait la réexpression d'un nombre restreint de gènes, ce qui donnerait a priori un risque de toxicité assez faible. Elle serait une alternative séduisante par rapport aux thérapies non sélectives déméthylant l'ADN ou inhibant l'action des désacétylases d'histones. On peut aussi émettre l'hypothèse que le ciblage de la protéine MBD2 puisse être utilisé en complément d'une de ces deux thérapies pour lesquelles les doses d'inhibiteurs seraient diminuées et donc moins susceptibles d'engendrer une toxicité.

Références bibliographiques

- Aapola, U., K. Kawasaki, et al. (2000). "Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family." Genomics **65**(3): 293-8.
- Achour, M., X. Jacq, et al. (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**(15): 2187-97.
- Aerts, S., G. Thijs, et al. (2004). "Comprehensive analysis of the base composition around the transcription start site in Metazoa." BMC Genomics **5**(1): 34.
- Allfrey, V. G., R. Faulkner, et al. (1964). "Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis." Proc Natl Acad Sci U S A **51**: 786-94.
- Allfrey, V. G., B. G. Pogo, et al. (1968). "Histone acetylation in insect chromosomes." Science **159**(812): 314-6.
- Amir, R. E., I. B. Van den Veyver, et al. (1999). "Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2." Nat Genet **23**(2): 185-8.
- Anderson, J. D. and J. Widom (2001). "Poly(dA-dT) promoter elements increase the equilibrium accessibility of nucleosomal DNA target sites." Mol Cell Biol **21**(11): 3830-9.
- Andres, M. E., C. Burger, et al. (1999). "CoREST: a functional corepressor required for regulation of neural-specific gene expression." Proc Natl Acad Sci U S A **96**(17): 9873-8.
- Angrisano, T., F. Lembo, et al. (2006). "TACC3 mediates the association of MBD2 with histone acetyltransferases and relieves transcriptional repression of methylated promoters." Nucleic Acids Res **34**(1): 364-72.
- Annunziato, A. T., M. B. Eason, et al. (1995). "Relationship between methylation and acetylation of arginine-rich histones in cycling and arrested HeLa cells." Biochemistry **34**(9): 2916-24.
- Antequera, F. and A. Bird (1993). "CpG islands." Exs **64**: 169-85.
- Antequera, F. and A. Bird (1993). "Number of CpG islands and genes in human and mouse." Proc Natl Acad Sci U S A **90**(24): 11995-9.
- Aranyi, T., B. A. Faucheux, et al. (2005). "The tissue-specific methylation of the human tyrosine hydroxylase gene reveals new regulatory elements in the first exon." J Neurochem **94**(1): 129-39.
- Arents, G., R. W. Burlingame, et al. (1991). "The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix." Proc Natl Acad Sci U S A **88**(22): 10148-52.
- Arima, Y., T. Hirota, et al. (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**(2): 131-42.

Auriol, E., L. M. Billard, et al. (2005). "Specific binding of the methyl binding domain protein 2 at the BRCA1-NBR2 locus." Nucleic Acids Res **33**(13): 4243-54.

Bachman, K. E., M. R. Rountree, et al. (2001). "Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin." J Biol Chem **276**(34): 32282-7.

Bacolla, A., S. Pradhan, et al. (1999). "Recombinant human DNA (cytosine-5) methyltransferase. II. Steady-state kinetics reveal allosteric activation by methylated dna." J Biol Chem **274**(46): 33011-9.

Badal, V., L. S. Chuang, et al. (2003). "CpG methylation of human papillomavirus type 16 DNA in cervical cancer cell lines and in clinical specimens: genomic hypomethylation correlates with carcinogenic progression." J Virol **77**(11): 6227-34.

Bai, S., K. Ghoshal, et al. (2005). "DNA methyltransferase 3b regulates nerve growth factor-induced differentiation of PC12 cells by recruiting histone deacetylase 2." Mol Cell Biol **25**(2): 751-66.

Bakker, J., X. Lin, et al. (2002). "Methyl-CpG binding domain protein 2 represses transcription from hypermethylated pi-class glutathione S-transferase gene promoters in hepatocellular carcinoma cells." J Biol Chem **277**(25): 22573-80.

Ballestar, E., M. F. Paz, et al. (2003). "Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer." Embo J **22**(23): 6335-45.

Balmer, D., J. Goldstine, et al. (2003). "Elevated methyl-CpG-binding protein 2 expression is acquired during postnatal human brain development and is correlated with alternative polyadenylation." J Mol Med **81**(1): 61-8.

Barski, A., S. Cuddapah, et al. (2007). "High-resolution profiling of histone methylations in the human genome." Cell **129**(4): 823-37.

Bartolomei, M. S. (2009). "Genomic imprinting: employing and avoiding epigenetic processes." Genes Dev **23**(18): 2124-33.

Baylin, S. B., J. G. Herman, et al. (1998). "Alterations in DNA methylation: a fundamental aspect of neoplasia." Adv Cancer Res **72**: 141-96.

Bednar, J., R. A. Horowitz, et al. (1998). "Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin." Proc Natl Acad Sci U S A **95**(24): 14173-8.

Bellacosa, A., L. Cicchillitti, et al. (1999). "MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1." Proc Natl Acad Sci U S A **96**(7): 3969-74.

Berger, J. and A. Bird (2005). "Role of MBD2 in gene regulation and tumorigenesis." Biochem Soc Trans **33**(Pt 6): 1537-40.

- Berger, S. L. (2007). "The complex language of chromatin regulation during transcription." Nature **447**(7143): 407-12.
- Berger, S. L., T. Kouzarides, et al. (2009). "An operational definition of epigenetics." Genes Dev **23**(7): 781-3.
- Bernstein, B. E., M. Kamal, et al. (2005). "Genomic maps and comparative analysis of histone modifications in human and mouse." Cell **120**(2): 169-81.
- Bestor, T., A. Laudano, et al. (1988). "Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases." J Mol Biol **203**(4): 971-83.
- Billard, L. M., F. Magdinier, et al. (2002). "MeCP2 and MBD2 expression during normal and pathological growth of the human mammary gland." Oncogene **21**(17): 2704-12.
- Bird, A. (2002). "DNA methylation patterns and epigenetic memory." Genes Dev **16**(1): 6-21.
- Bird, A. (2007). "Perceptions of epigenetics." Nature **447**(7143): 396-8.
- Bird, A., M. Taggart, et al. (1985). "A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA." Cell **40**(1): 91-9.
- Boeger, H., J. Griesenbeck, et al. (2003). "Nucleosomes unfold completely at a transcriptionally active promoter." Mol Cell **11**(6): 1587-98.
- Boeke, J., O. Ammerpohl, et al. (2000). "The minimal repression domain of MBD2b overlaps with the methyl-CpG-binding domain and binds directly to Sin3A." J Biol Chem **275**(45): 34963-7.
- Bonenfant, D., H. Towbin, et al. (2007). "Analysis of dynamic changes in post-translational modifications of human histones during cell cycle by mass spectrometry." Mol Cell Proteomics **6**(11): 1917-32.
- Bonfils, C., N. Beaulieu, et al. (2000). "Characterization of the human DNA methyltransferase splice variant Dnmt1b." J Biol Chem **275**(15): 10754-60.
- Bostick, M., J. K. Kim, et al. (2007). "UHRF1 plays a role in maintaining DNA methylation in mammalian cells." Science **317**(5845): 1760-4.
- Bourc'his, D., G. L. Xu, et al. (2001). "Dnmt3L and the establishment of maternal genomic imprints." Science **294**(5551): 2536-9.
- Bracaglia, G., B. Conca, et al. (2009). "Methyl-CpG-binding protein 2 is phosphorylated by homeodomain-interacting protein kinase 2 and contributes to apoptosis." EMBO Rep **10**(12): 1327-33.
- Brown, M. A., R. J. Sims, 3rd, et al. (2006). "Identification and characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-specific methyltransferase that interacts with the Sin3 histone deacetylase complex." Mol Cancer **5**: 26.

- Burgers, W. A., L. Blanchon, et al. (2007). "Viral oncoproteins target the DNA methyltransferases." *Oncogene* **26**(11): 1650-5.
- Cameron, E. E., K. E. Bachman, et al. (1999). "Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer." *Nat Genet* **21**(1): 103-7.
- Campbell, P. M., V. Bovenzi, et al. (2004). "Methylated DNA-binding protein 2 antisense inhibitors suppress tumorigenesis of human cancer cell lines in vitro and in vivo." *Carcinogenesis* **25**(4): 499-507.
- Candau, R., P. A. Moore, et al. (1996). "Identification of human proteins functionally conserved with the yeast putative adaptors ADA2 and GCN5." *Mol Cell Biol* **16**(2): 593-602.
- Cassel, S., M. O. Revel, et al. (2004). "Expression of the methyl-CpG-binding protein MeCP2 in rat brain. An ontogenetic study." *Neurobiol Dis* **15**(2): 206-11.
- Chahrour, M., S. Y. Jung, et al. (2008). "MeCP2, a key contributor to neurological disease, activates and represses transcription." *Science* **320**(5880): 1224-9.
- Chatagnon, A., E. Ballestar, et al. "A role for methyl-CpG binding domain protein 2 in the modulation of the estrogen response of pS2/TFF1 gene." *PLoS One* **5**(3): e9665.
- Chatagnon, A., S. Bougel, et al. (2009). "Specific association between the methyl-CpG-binding domain protein 2 and the hypermethylated region of the human telomerase reverse transcriptase promoter in cancer cells." *Carcinogenesis* **30**(1): 28-34.
- Chen, T., Y. Ueda, et al. (2002). "A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation." *J Biol Chem* **277**(41): 38746-54.
- Chen, W. G., Q. Chang, et al. (2003). "Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2." *Science* **302**(5646): 885-9.
- Chen, W. Y., X. Zeng, et al. (2003). "Heterozygous disruption of Hic1 predisposes mice to a gender-dependent spectrum of malignant tumors." *Nat Genet* **33**(2): 197-202.
- Cheng, P., C. Schmutte, et al. (1997). "Alterations in DNA methylation are early, but not initial, events in ovarian tumorigenesis." *Br J Cancer* **75**(3): 396-402.
- Cheng, X. (1995). "Structure and function of DNA methyltransferases." *Annu Rev Biophys Biomol Struct* **24**: 293-318.
- Chuang, L. S., H. I. Ian, et al. (1997). "Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1." *Science* **277**(5334): 1996-2000.
- Citterio, E., R. Papait, et al. (2004). "Np95 is a histone-binding protein endowed with ubiquitin ligase activity." *Mol Cell Biol* **24**(6): 2526-35.

- Clouaire, T., J. I. de Las Heras, et al. "Recruitment of MBD1 to target genes requires sequence-specific interaction of the MBD domain with methylated DNA." Nucleic Acids Res.
- Coolen, M. W., C. Stirzaker, et al. "Consolidation of the cancer genome into domains of repressive chromatin by long-range epigenetic silencing (LRES) reduces transcriptional plasticity." Nat Cell Biol **12**(3): 235-46.
- Cortellino, S., D. Turner, et al. (2003). "The base excision repair enzyme MED1 mediates DNA damage response to antitumor drugs and is associated with mismatch repair system integrity." Proc Natl Acad Sci U S A **100**(25): 15071-6.
- Costa, F. F., E. A. Seftor, et al. (2009). "Epigenetically reprogramming metastatic tumor cells with an embryonic microenvironment." Epigenomics **1**(2): 387-398.
- Covic, M., P. O. Hassa, et al. (2005). "Arginine methyltransferase CARM1 is a promoter-specific regulator of NF-kappaB-dependent gene expression." Embo J **24**(1): 85-96.
- Cross, S. H., J. A. Charlton, et al. (1994). "Purification of CpG islands using a methylated DNA binding column." Nat Genet **6**(3): 236-44.
- Curradi, M., A. Izzo, et al. (2002). "Molecular mechanisms of gene silencing mediated by DNA methylation." Mol Cell Biol **22**(9): 3157-73.
- Dahl, J. A., A. H. Reiner, et al. "Histone H3 lysine 27 methylation asymmetry on developmentally-regulated promoters distinguish the first two lineages in mouse preimplantation embryos." PLoS One **5**(2): e9150.
- Dai, S. D., Y. Wang, et al. "Kaiso is expressed in lung cancer: its expression and localization is affected by p120ctn." Lung Cancer **67**(2): 205-15.
- Daniel, J. M. and A. B. Reynolds (1999). "The catenin p120(ctn) interacts with Kaiso, a novel BTB/POZ domain zinc finger transcription factor." Mol Cell Biol **19**(5): 3614-23.
- Daniel, J. M., C. M. Spring, et al. (2002). "The p120(ctn)-binding partner Kaiso is a bi-modal DNA-binding protein that recognizes both a sequence-specific consensus and methylated CpG dinucleotides." Nucleic Acids Res **30**(13): 2911-9.
- Davalos, V. and M. Esteller "MicroRNAs and cancer epigenetics: a macroevolution." Curr Opin Oncol **22**(1): 35-45.
- Davey, C. A. and T. J. Richmond (2002). "DNA-dependent divalent cation binding in the nucleosome core particle." Proc Natl Acad Sci U S A **99**(17): 11169-74.
- David, G., M. A. Neptune, et al. (2002). "SUMO-1 modification of histone deacetylase 1 (HDAC1) modulates its biological activities." J Biol Chem **277**(26): 23658-63.
- De La Rosa-Velazquez, I. A., H. Rincon-Arango, et al. (2007). "Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF." Cancer Res **67**(6): 2577-85.

- De Marzo, A. M., V. L. Marchi, et al. (1999). "Abnormal regulation of DNA methyltransferase expression during colorectal carcinogenesis." Cancer Res **59**(16): 3855-60.
- De Smet, C., C. Lurquin, et al. (1999). "DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter." Mol Cell Biol **19**(11): 7327-35.
- Defossez, P. A., K. F. Kelly, et al. (2005). "The human enhancer blocker CTC-binding factor interacts with the transcription factor Kaiso." J Biol Chem **280**(52): 43017-23.
- Della Ragione, F., A. Tiunova, et al. (2006). "The X-linked methyl binding protein gene Kaiso is highly expressed in brain but is not mutated in Rett syndrome patients." Gene **373**: 83-9.
- Detich, N., J. Theberge, et al. (2002). "Promoter-specific activation and demethylation by MBD2/demethylase." J Biol Chem **277**(39): 35791-4.
- Dhordain, P., O. Albagli, et al. (2000). "Colocalization and heteromerization between the two human oncogene POZ/zinc finger proteins, LAZ3 (BCL6) and PLZF." Oncogene **19**(54): 6240-50.
- Dickerson, R. E., D. S. Goodsell, et al. (1994). ""...the tyranny of the lattice..."" Proc Natl Acad Sci U S A **91**(9): 3579-83.
- Doherty, A. S., M. S. Bartolomei, et al. (2002). "Regulation of stage-specific nuclear translocation of Dnmt1o during preimplantation mouse development." Dev Biol **242**(2): 255-66.
- Donaldson, N. S., C. L. Nordgaard, et al. "Kaiso regulates Znf131-mediated transcriptional activation." Exp Cell Res **316**(10): 1692-705.
- Dragich, J. M., Y. H. Kim, et al. (2007). "Differential distribution of the MeCP2 splice variants in the postnatal mouse brain." J Comp Neurol **501**(4): 526-42.
- Dvash, T., N. Lavon, et al. "Variations of X chromosome inactivation occur in early passages of female human embryonic stem cells." PLoS One **5**(6): e11330.
- Eads, C. A., K. D. Danenberg, et al. (1999). "CpG island hypermethylation in human colorectal tumors is not associated with DNA methyltransferase overexpression." Cancer Res **59**(10): 2302-6.
- Eberharter, A. and P. B. Becker (2002). "Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics." EMBO Rep **3**(3): 224-9.
- Eden, A., F. Gaudet, et al. (2003). "Chromosomal instability and tumors promoted by DNA hypomethylation." Science **300**(5618): 455.
- Ehrlich, M., M. A. Gama-Sosa, et al. (1982). "Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells." Nucleic Acids Res **10**(8): 2709-21.

- Ehrlich, M., C. B. Woods, et al. (2006). "Quantitative analysis of associations between DNA hypermethylation, hypomethylation, and DNMT RNA levels in ovarian tumors." Oncogene **25**(18): 2636-45.
- Elenbaas, B., L. Spirio, et al. (2001). "Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells." Genes Dev **15**(1): 50-65.
- Ellgaard, E. G. (1967). "Gene activation without histone acetylation in *Drosophila melanogaster*." Science **157**(792): 1070-2.
- Eskeland, R., M. Leeb, et al. "Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination." Mol Cell **38**(3): 452-64.
- Esteller, M., S. Tortola, et al. (2000). "Hypermethylation-associated inactivation of p14(ARF) is independent of p16(INK4a) methylation and p53 mutational status." Cancer Res **60**(1): 129-33.
- Esteve, P. O., H. G. Chin, et al. (2006). "Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication." Genes Dev **20**(22): 3089-103.
- Evans, P. M., W. Zhang, et al. (2007). "Kruppel-like factor 4 is acetylated by p300 and regulates gene transcription via modulation of histone acetylation." J Biol Chem **282**(47): 33994-4002.
- Fang, J. Y., Z. H. Cheng, et al. (2004). "Expression of Dnmt1, demethylase, MeCP2 and methylation of tumor-related genes in human gastric cancer." World J Gastroenterol **10**(23): 3394-8.
- Feinberg, A. P., C. W. Gehrke, et al. (1988). "Reduced genomic 5-methylcytosine content in human colonic neoplasia." Cancer Res **48**(5): 1159-61.
- Felsenfeld, G. and M. Groudine (2003). "Controlling the double helix." Nature **421**(6921): 448-53.
- Feng, Q., B. He, et al. (2009). "Biochemical control of CARM1 enzymatic activity by phosphorylation." J Biol Chem **284**(52): 36167-74.
- Feng, Q., H. Wang, et al. (2002). "Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain." Curr Biol **12**(12): 1052-8.
- Feng, Q. and Y. Zhang (2001). "The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes." Genes Dev **15**(7): 827-32.
- Filion, G. J., S. Zhenilo, et al. (2006). "A family of human zinc finger proteins that bind methylated DNA and repress transcription." Mol Cell Biol **26**(1): 169-81.
- Finnegan, E. J., W. J. Peacock, et al. (2000). "DNA methylation, a key regulator of plant development and other processes." Curr Opin Genet Dev **10**(2): 217-23.

- Fleisher, A. S., M. Esteller, et al. (2001). "Hypermethylation of the hMLH1 gene promoter is associated with microsatellite instability in early human gastric neoplasia." Oncogene **20**(3): 329-35.
- Fraga, M. F., E. Ballestar, et al. (2003). "The affinity of different MBD proteins for a specific methylated locus depends on their intrinsic binding properties." Nucleic Acids Res **31**(6): 1765-74.
- Franklin, S. G. and A. Zweidler (1977). "Non-allelic variants of histones 2a, 2b and 3 in mammals." Nature **266**(5599): 273-5.
- Fujita, H., R. Fujii, et al. (2003). "Antithetic effects of MBD2a on gene regulation." Mol Cell Biol **23**(8): 2645-57.
- Fujita, N., S. Takebayashi, et al. (1999). "Methylation-mediated transcriptional silencing in euchromatin by methyl-CpG binding protein MBD1 isoforms." Mol Cell Biol **19**(9): 6415-26.
- Fujita, N., S. Watanabe, et al. (2003). "Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression." J Biol Chem **278**(26): 24132-8.
- Fuks, F., W. A. Burgers, et al. (2000). "DNA methyltransferase Dnmt1 associates with histone deacetylase activity." Nat Genet **24**(1): 88-91.
- Fuks, F., W. A. Burgers, et al. (2001). "Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription." Embo J **20**(10): 2536-44.
- Fuks, F., P. J. Hurd, et al. (2003). "The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation." J Biol Chem **278**(6): 4035-40.
- Fullwood, M. J., M. H. Liu, et al. (2009). "An oestrogen-receptor-alpha-bound human chromatin interactome." Nature **462**(7269): 58-64.
- Gallwitz, D. and C. E. Sekeris (1969). "Stimulation of RNA polymerase activity of rat liver nuclei by cortisol in vitro independent of effects on the acetylation and methylation of histones." FEBS Lett **3**(2): 99-102.
- Gama-Sosa, M. A., V. A. Slagel, et al. (1983). "The 5-methylcytosine content of DNA from human tumors." Nucleic Acids Res **11**(19): 6883-94.
- Gardiner-Garden, M. and M. Frommer (1987). "CpG islands in vertebrate genomes." J Mol Biol **196**(2): 261-82.
- Gonzalo, S., M. Garcia-Cao, et al. (2005). "Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin." Nat Cell Biol **7**(4): 420-8.
- Grunstein, M. (1997). "Histone acetylation in chromatin structure and transcription." Nature **389**(6649): 349-52.
- Gu, W., X. L. Shi, et al. (1997). "Synergistic activation of transcription by CBP and p53." Nature **387**(6635): 819-23.

- Guy, J., B. Hendrich, et al. (2001). "A mouse *Mecp2*-null mutation causes neurological symptoms that mimic Rett syndrome." *Nat Genet* **27**(3): 322-6.
- Hamamoto, R., Y. Furukawa, et al. (2004). "SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells." *Nat Cell Biol* **6**(8): 731-40.
- Harikrishnan, K. N., M. Z. Chow, et al. (2005). "Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing." *Nat Genet* **37**(3): 254-64.
- Hata, K., M. Okano, et al. (2002). "Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice." *Development* **129**(8): 1983-93.
- Hattori, N., T. Abe, et al. (2004). "Preference of DNA methyltransferases for CpG islands in mouse embryonic stem cells." *Genome Res* **14**(9): 1733-40.
- Heintzman, N. D., R. K. Stuart, et al. (2007). "Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome." *Nat Genet* **39**(3): 311-8.
- Hendrich, B. and A. Bird (1998). "Identification and characterization of a family of mammalian methyl-CpG binding proteins." *Mol Cell Biol* **18**(11): 6538-47.
- Hendrich, B., J. Guy, et al. (2001). "Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development." *Genes Dev* **15**(6): 710-23.
- Herrera, J. E., K. L. West, et al. (2000). "Histone H1 is a specific repressor of core histone acetylation in chromatin." *Mol Cell Biol* **20**(2): 523-9.
- Hildebrand, J. D. and P. Soriano (2002). "Overlapping and unique roles for C-terminal binding protein 1 (CtBP1) and CtBP2 during mouse development." *Mol Cell Biol* **22**(15): 5296-307.
- Hodge, D. R., B. Peng, et al. (2005). "Interleukin 6 supports the maintenance of p53 tumor suppressor gene promoter methylation." *Cancer Res* **65**(11): 4673-82.
- Hodge, D. R., W. Xiao, et al. (2001). "Interleukin-6 regulation of the human DNA methyltransferase (HDNMT) gene in human erythroleukemia cells." *J Biol Chem* **276**(43): 39508-11.
- Hodges, E., A. D. Smith, et al. (2009). "High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing." *Genome Res* **19**(9): 1593-605.
- Hong, L., G. P. Schroth, et al. (1993). "Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA." *J Biol Chem* **268**(1): 305-14.
- Hong, S., A. Derfoul, et al. (2009). "A novel domain in histone deacetylase 1 and 2 mediates repression of cartilage-specific genes in human chondrocytes." *Faseb J* **23**(10): 3539-52.

- Hong, W., M. Nakazawa, et al. (2005). "FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1." Embo J **24**(13): 2367-78.
- Horike, S., S. Cai, et al. (2005). "Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome." Nat Genet **37**(1): 31-40.
- Hsieh, C. L. (1999). "In vivo activity of murine de novo methyltransferases, Dnmt3a and Dnmt3b." Mol Cell Biol **19**(12): 8211-8.
- Hu, K., X. Nan, et al. (2006). "Testing for association between MeCP2 and the brahma-associated SWI/SNF chromatin-remodeling complex." Nat Genet **38**(9): 962-4; author reply 964-7.
- Huntriss, J., M. Hinkins, et al. (2004). "Expression of mRNAs for DNA methyltransferases and methyl-CpG-binding proteins in the human female germ line, preimplantation embryos, and embryonic stem cells." Mol Reprod Dev **67**(3): 323-36.
- Hutchins, A. S., A. C. Mullen, et al. (2002). "Gene silencing quantitatively controls the function of a developmental trans-activator." Mol Cell **10**(1): 81-91.
- Iida, T., I. Suetake, et al. (2002). "PCNA clamp facilitates action of DNA cytosine methyltransferase 1 on hemimethylated DNA." Genes Cells **7**(10): 997-1007.
- Illingworth, R., A. Kerr, et al. (2008). "A novel CpG island set identifies tissue-specific methylation at developmental gene loci." PLoS Biol **6**(1): e22.
- Ioshikhes, I. P. and M. Q. Zhang (2000). "Large-scale human promoter mapping using CpG islands." Nat Genet **26**(1): 61-3.
- Isono, K., K. Nemoto, et al. (2006). "Overlapping roles for homeodomain-interacting protein kinases hipk1 and hipk2 in the mediation of cell growth in response to morphogenetic and genotoxic signals." Mol Cell Biol **26**(7): 2758-71.
- Issa, J. P. (2004). "CpG island methylator phenotype in cancer." Nat Rev Cancer **4**(12): 988-93.
- Ivanov, M. A., B. Lamrihi, et al. (2003). "Enhanced antitumor activity of a combination of MBD2-antisense electrotransfer gene therapy and bleomycin electrochemotherapy." J Gene Med **5**(10): 893-9.
- Jaenisch, R. and A. Bird (2003). "Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals." Nat Genet **33** **Suppl**: 245-54.
- Jeanblanc, M., M. Mousli, et al. (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**(49): 7337-45.
- Jeffery, L. and S. Nakielny (2004). "Components of the DNA methylation system of chromatin control are RNA-binding proteins." J Biol Chem **279**(47): 49479-87.

Jeppesen, P. and B. M. Turner (1993). "The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression." Cell **74**(2): 281-9.

Jiang, C. L., S. G. Jin, et al. (2004). "MBD3L1 is a transcriptional repressor that interacts with methyl-CpG-binding protein 2 (MBD2) and components of the NuRD complex." J Biol Chem **279**(50): 52456-64.

Jin, S. G., C. L. Jiang, et al. (2005). "MBD3L2 interacts with MBD3 and components of the NuRD complex and can oppose MBD2-MeCP1-mediated methylation silencing." J Biol Chem **280**(13): 12700-9.

Johns, E. W., D. M. Phillips, et al. (1961). "The electrophoresis of histones and histone fractions on starch gel." Biochem J **80**: 189-93.

Johnson, L. M., J. A. Law, et al. (2008). "SRA-domain proteins required for DRM2-mediated de novo DNA methylation." PLoS Genet **4**(11): e1000280.

Jones, P. L., G. J. Veenstra, et al. (1998). "Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription." Nat Genet **19**(2): 187-91.

Jorgensen, H. F., K. Adie, et al. (2006). "Engineering a high-affinity methyl-CpG-binding protein." Nucleic Acids Res **34**(13): e96.

Jorgensen, H. F., I. Ben-Porath, et al. (2004). "Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains." Mol Cell Biol **24**(8): 3387-95.

Kaludov, N. K. and A. P. Wolffe (2000). "MeCP2 driven transcriptional repression in vitro: selectivity for methylated DNA, action at a distance and contacts with the basal transcription machinery." Nucleic Acids Res **28**(9): 1921-8.

Kanai, Y. and S. Hirohashi (2007). "Alterations of DNA methylation associated with abnormalities of DNA methyltransferases in human cancers during transition from a precancerous to a malignant state." Carcinogenesis **28**(12): 2434-42.

Kanai, Y., S. Ushijima, et al. (1999). "Reduced mRNA expression of the DNA demethylase, MBD2, in human colorectal and stomach cancers." Biochem Biophys Res Commun **264**(3): 962-6.

Kantor, B., K. Makedonski, et al. (2003). "Expression and localization of components of the histone deacetylases multiprotein repressory complexes in the mouse preimplantation embryo." Gene Expr Patterns **3**(6): 697-702.

Karagianni, P., L. Amazit, et al. (2008). "ICBP90, a novel methyl K9 H3 binding protein linking protein ubiquitination with heterochromatin formation." Mol Cell Biol **28**(2): 705-17.

Kelly, K. F., A. A. Otchere, et al. (2004). "Nuclear import of the BTB/POZ transcriptional regulator Kaiso." J Cell Sci **117**(Pt 25): 6143-52.

- Kelly, K. F., C. M. Spring, et al. (2004). "NLS-dependent nuclear localization of p120ctn is necessary to relieve Kaiso-mediated transcriptional repression." J Cell Sci **117**(Pt 13): 2675-86.
- Keshet, I., J. Lieman-Hurwitz, et al. (1986). "DNA methylation affects the formation of active chromatin." Cell **44**(4): 535-43.
- Khulan, B., R. F. Thompson, et al. (2006). "Comparative isoschizomer profiling of cytosine methylation: the HELP assay." Genome Res **16**(8): 1046-55.
- Kiefer, H., F. Chatail-Hermitte, et al. (2005). "ZENON, a novel POZ Kruppel-like DNA binding protein associated with differentiation and/or survival of late postmitotic neurons." Mol Cell Biol **25**(5): 1713-29.
- Kim, G. D., J. Ni, et al. (2002). "Co-operation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases." Embo J **21**(15): 4183-95.
- Kim, J. K., P. O. Esteve, et al. (2009). "UHRF1 binds G9a and participates in p21 transcriptional regulation in mammalian cells." Nucleic Acids Res **37**(2): 493-505.
- Kim, J. S., J. Han, et al. (2005). "Aberrant methylation of H-cadherin (CDH13) promoter is associated with tumor progression in primary nonsmall cell lung carcinoma." Cancer **104**(9): 1825-33.
- Kim, S. M., H. J. Kee, et al. (2006). "Characterization of a novel WHSC1-associated SET domain protein with H3K4 and H3K27 methyltransferase activity." Biochem Biophys Res Commun **345**(1): 318-23.
- Kim, S. W., J. I. Park, et al. (2004). "Non-canonical Wnt signals are modulated by the Kaiso transcriptional repressor and p120-catenin." Nat Cell Biol **6**(12): 1212-20.
- Kim, Y. I., A. Giuliano, et al. (1994). "Global DNA hypomethylation increases progressively in cervical dysplasia and carcinoma." Cancer **74**(3): 893-9.
- Kimura, H. and K. Shiota (2003). "Methyl-CpG-binding protein, MeCP2, is a target molecule for maintenance DNA methyltransferase, Dnmt1." J Biol Chem **278**(7): 4806-12.
- Kishi, N. and J. D. Macklis (2004). "MECP2 is progressively expressed in post-migratory neurons and is involved in neuronal maturation rather than cell fate decisions." Mol Cell Neurosci **27**(3): 306-21.
- Kishikawa, S., T. Murata, et al. (2002). "Regulation of transcription of the Dnmt1 gene by Sp1 and Sp3 zinc finger proteins." Eur J Biochem **269**(12): 2961-70.
- Kishikawa, S., T. Murata, et al. (2003). "Control elements of Dnmt1 gene are regulated in cell-cycle dependent manner." Nucleic Acids Res Suppl(3): 307-8.
- Klose, R. J., S. A. Sarraf, et al. (2005). "DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG." Mol Cell **19**(5): 667-78.

- Klose, R. J., K. Yamane, et al. (2006). "The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36." Nature **442**(7100): 312-6.
- Klug, A., D. Rhodes, et al. (1980). "A low resolution structure for the histone core of the nucleosome." Nature **287**(5782): 509-16.
- Kn, H., S. Bassal, et al. (2004). "Expression analysis of the epigenetic methyltransferases and methyl-CpG binding protein families in the normal B-cell and B-cell chronic lymphocytic leukemia (CLL)." Cancer Biol Ther **3**(10): 989-94.
- Kokura, K., S. C. Kaul, et al. (2001). "The Ski protein family is required for MeCP2-mediated transcriptional repression." J Biol Chem **276**(36): 34115-21.
- Kondo, E., Z. Gu, et al. (2005). "The thymine DNA glycosylase MBD4 represses transcription and is associated with methylated p16(INK4a) and hMLH1 genes." Mol Cell Biol **25**(11): 4388-96.
- Kornberg, R. D. and J. O. Thomas (1974). "Chromatin structure; oligomers of the histones." Science **184**(139): 865-8.
- Kouzarides, T. (2007). "Chromatin modifications and their function." Cell **128**(4): 693-705.
- Kouzminova, E. and E. U. Selker (2001). "dim-2 encodes a DNA methyltransferase responsible for all known cytosine methylation in Neurospora." Embo J **20**(15): 4309-23.
- Lachner, M., D. O'Carroll, et al. (2001). "Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins." Nature **410**(6824): 116-20.
- Lane, N., W. Dean, et al. (2003). "Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse." Genesis **35**(2): 88-93.
- Larsen, F., G. Gundersen, et al. (1992). "CpG islands as gene markers in the human genome." Genomics **13**(4): 1095-107.
- Lauster, R., T. A. Trautner, et al. (1989). "Cytosine-specific type II DNA methyltransferases. A conserved enzyme core with variable target-recognizing domains." J Mol Biol **206**(2): 305-12.
- Le Guezennec, X., M. Vermeulen, et al. (2006). "MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties." Mol Cell Biol **26**(3): 843-51.
- Lee, J., S. J. Jang, et al. "Presence of 5-methylcytosine in CpNpG trinucleotides in the human genome." Genomics.
- Lehnertz, B., Y. Ueda, et al. (2003). "Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin." Curr Biol **13**(14): 1192-200.
- Lei, H., S. P. Oh, et al. (1996). "De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells." Development **122**(10): 3195-205.

- Lembo, F., R. Pero, et al. (2003). "MBDin, a novel MBD2-interacting protein, relieves MBD2 repression potential and reactivates transcription from methylated promoters." Mol Cell Biol **23**(5): 1656-65.
- Lengauer, C., K. W. Kinzler, et al. (1997). "DNA methylation and genetic instability in colorectal cancer cells." Proc Natl Acad Sci U S A **94**(6): 2545-50.
- Leonhardt, H., A. W. Page, et al. (1992). "A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei." Cell **71**(5): 865-73.
- Leu, Y. W., F. Rahmatpanah, et al. (2003). "Double RNA interference of DNMT3b and DNMT1 enhances DNA demethylation and gene reactivation." Cancer Res **63**(19): 6110-5.
- Liang, G., J. C. Lin, et al. (2004). "Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome." Proc Natl Acad Sci U S A **101**(19): 7357-62.
- Lin, X. and W. G. Nelson (2003). "Methyl-CpG-binding domain protein-2 mediates transcriptional repression associated with hypermethylated GSTP1 CpG islands in MCF-7 breast cancer cells." Cancer Res **63**(2): 498-504.
- Lin, X., M. Tascilar, et al. (2001). "GSTP1 CpG island hypermethylation is responsible for the absence of GSTP1 expression in human prostate cancer cells." Am J Pathol **159**(5): 1815-26.
- Liu, C., Z. Q. Teng, et al. "Epigenetic regulation of miR-184 by MBD1 governs neural stem cell proliferation and differentiation." Cell Stem Cell **6**(5): 433-44.
- Lo Coco, F., B. H. Ye, et al. (1994). "Rearrangements of the BCL6 gene in diffuse large cell non-Hodgkin's lymphoma." Blood **83**(7): 1757-9.
- Lopes, E. C., E. Valls, et al. (2008). "Kaiso contributes to DNA methylation-dependent silencing of tumor suppressor genes in colon cancer cell lines." Cancer Res **68**(18): 7258-63.
- Lopez-Serra, L., E. Ballestar, et al. (2006). "A profile of methyl-CpG binding domain protein occupancy of hypermethylated promoter CpG islands of tumor suppressor genes in human cancer." Cancer Res **66**(17): 8342-6.
- Lopez-Serra, L., E. Ballestar, et al. (2008). "Unmasking of epigenetically silenced candidate tumor suppressor genes by removal of methyl-CpG-binding domain proteins." Oncogene **27**(25): 3556-66.
- Luger, K., A. W. Mader, et al. (1997). "Crystal structure of the nucleosome core particle at 2.8 Å resolution." Nature **389**(6648): 251-60.
- Lunyak, V. V., R. Burgess, et al. (2002). "Corepressor-dependent silencing of chromosomal regions encoding neuronal genes." Science **298**(5599): 1747-52.

- Luo, S. W., C. Zhang, et al. (2009). "Regulation of heterochromatin remodelling and myogenin expression during muscle differentiation by FAK interaction with MBD2." Embo J **28**(17): 2568-82.
- Luo, Y., W. Jian, et al. (2009). "Trans-regulation of histone deacetylase activities through acetylation." J Biol Chem **284**(50): 34901-10.
- Lyko, F., B. H. Ramsahoye, et al. (2000). "DNA methylation in *Drosophila melanogaster*." Nature **408**(6812): 538-40.
- Lyst, M. J., X. Nan, et al. (2006). "Regulation of MBD1-mediated transcriptional repression by SUMO and PIAS proteins." Embo J **25**(22): 5317-28.
- Macdonald, J. L., A. Verster, et al. "MBD2 and MeCP2 regulate distinct transitions in the stage-specific differentiation of olfactory receptor neurons." Mol Cell Neurosci **44**(1): 55-67.
- Malone, C. S., M. D. Miner, et al. (2001). "CmC(A/T)GG DNA methylation in mature B cell lymphoma gene silencing." Proc Natl Acad Sci U S A **98**(18): 10404-9.
- Margot, J. B., M. C. Cardoso, et al. (2001). "Mammalian DNA methyltransferases show different subnuclear distributions." J Cell Biochem **83**(3): 373-9.
- Martin Caballero, I., J. Hansen, et al. (2009). "The methyl-CpG binding proteins MeCP2, Mbd2 and Kaiso are dispensable for mouse embryogenesis, but play a redundant function in neural differentiation." PLoS One **4**(1): e4315.
- Martinowich, K., D. Hattori, et al. (2003). "DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation." Science **302**(5646): 890-3.
- Masson, D., N. Rioux-Leclercq, et al. "Loss of expression of TIMP3 in clear cell renal cell carcinoma." Eur J Cancer **46**(8): 1430-7.
- Matarazzo, M. R., M. L. De Bonis, et al. (2007). "Multiple binding of methyl-CpG and polycomb proteins in long-term gene silencing events." J Cell Physiol **210**(3): 711-9.
- Matsuda, R., T. Hori, et al. "Identification and characterization of the two isoforms of the vertebrate H2A.Z histone variant." Nucleic Acids Res.
- Mayer, W., A. Niveleau, et al. (2000). "Demethylation of the zygotic paternal genome." Nature **403**(6769): 501-2.
- Meehan, R. R., J. D. Lewis, et al. (1992). "Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA." Nucleic Acids Res **20**(19): 5085-92.
- Mertineit, C., J. A. Yoder, et al. (1998). "Sex-specific exons control DNA methyltransferase in mammalian germ cells." Development **125**(5): 889-97.
- Millar, C. B., J. Guy, et al. (2002). "Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice." Science **297**(5580): 403-5.

- Mizuno, S., T. Chijiwa, et al. (2001). "Expression of DNA methyltransferases DNMT1, 3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia." Blood **97**(5): 1172-9.
- Mohandas, T., R. S. Sparkes, et al. (1981). "Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation." Science **211**(4480): 393-6.
- Morel, A. P., M. Lievre, et al. (2008). "Generation of breast cancer stem cells through epithelial-mesenchymal transition." PLoS One **3**(8): e2888.
- Morris, S. A., B. Rao, et al. (2007). "Identification of histone H3 lysine 36 acetylation as a highly conserved histone modification." J Biol Chem **282**(10): 7632-40.
- Muller-Tidow, C., K. Kugler, et al. (2001). "Loss of expression of HDAC-recruiting methyl-CpG-binding domain proteins in human cancer." Br J Cancer **85**(8): 1168-74.
- Muller, H. M., H. Fiegl, et al. (2003). "MeCP2 and MBD2 expression in human neoplastic and non-neoplastic breast tissue and its association with oestrogen receptor status." Br J Cancer **89**(10): 1934-9.
- Muller, I., F. Wischnewski, et al. "Promoter- and cell-specific epigenetic regulation of CD44, Cyclin D2, GLIPR1 and PTEN by Methyl-CpG binding proteins and histone modifications." BMC Cancer **10**(1): 297.
- Muscat, G. E., L. J. Burke, et al. (1998). "The corepressor N-CoR and its variants RIP13a and RIP13Delta1 directly interact with the basal transcription factors TFIIB, TAFII32 and TAFII70." Nucleic Acids Res **26**(12): 2899-907.
- Na, J., K. Lykke-Andersen, et al. (2007). "Dishevelled proteins regulate cell adhesion in mouse blastocyst and serve to monitor changes in Wnt signaling." Dev Biol **302**(1): 40-9.
- Nakayama, J., J. C. Rice, et al. (2001). "Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly." Science **292**(5514): 110-3.
- Nan, X., F. J. Campoy, et al. (1997). "MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin." Cell **88**(4): 471-81.
- Nan, X., H. H. Ng, et al. (1998). "Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex." Nature **393**(6683): 386-9.
- Nan, X., P. Tate, et al. (1996). "DNA methylation specifies chromosomal localization of MeCP2." Mol Cell Biol **16**(1): 414-21.
- Narayan, A., W. Ji, et al. (1998). "Hypomethylation of pericentromeric DNA in breast adenocarcinomas." Int J Cancer **77**(6): 833-8.
- Naruse, Y., T. Aoki, et al. (1999). "Neural restrictive silencer factor recruits mSin3 and histone deacetylase complex to repress neuron-specific target genes." Proc Natl Acad Sci U S A **96**(24): 13691-6.
- Ng, H. H., P. Jeppesen, et al. (2000). "Active repression of methylated genes by the chromosomal protein MBD1." Mol Cell Biol **20**(4): 1394-406.

- Ng, H. H., Y. Zhang, et al. (1999). "MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex." Nat Genet **23**(1): 58-61.
- Nicolas, E., C. Roumillac, et al. (2003). "Balance between acetylation and methylation of histone H3 lysine 9 on the E2F-responsive dihydrofolate reductase promoter." Mol Cell Biol **23**(5): 1614-22.
- Nielsen, S. J., R. Schneider, et al. (2001). "Rb targets histone H3 methylation and HP1 to promoters." Nature **412**(6846): 561-5.
- Nikitina, T., R. P. Ghosh, et al. (2007). "MeCP2-chromatin interactions include the formation of chromosome-like structures and are altered in mutations causing Rett syndrome." J Biol Chem **282**(38): 28237-45.
- Noh, E. J., D. S. Lim, et al. (2009). "A novel role for methyl CpG-binding domain protein 3, a component of the histone deacetylase complex, in regulation of cell cycle progression and cell death." Biochem Biophys Res Commun **378**(3): 332-7.
- Noll, M. (1974). "Internal structure of the chromatin subunit." Nucleic Acids Res **1**(11): 1573-8.
- Oikawa, Y., E. Matsuda, et al. (2008). "Down-regulation of CIBZ, a novel substrate of caspase-3, induces apoptosis." J Biol Chem **283**(21): 14242-7.
- Okano, M., D. W. Bell, et al. (1999). "DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development." Cell **99**(3): 247-57.
- Okano, M., S. Xie, et al. (1998). "Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases." Nat Genet **19**(3): 219-20.
- Okano, M., S. Xie, et al. (1998). "Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells." Nucleic Acids Res **26**(11): 2536-40.
- Ostler, K. R., E. M. Davis, et al. (2007). "Cancer cells express aberrant DNMT3B transcripts encoding truncated proteins." Oncogene **26**(38): 5553-63.
- Oswald, J., S. Engemann, et al. (2000). "Active demethylation of the paternal genome in the mouse zygote." Curr Biol **10**(8): 475-8.
- Oudet, P., M. Gross-Bellard, et al. (1975). "Electron microscopic and biochemical evidence that chromatin structure is a repeating unit." Cell **4**(4): 281-300.
- Oue, N., K. Kuraoka, et al. (2001). "DNA methylation status of hMLH1, p16(INK4a), and CDH1 is not associated with mRNA expression levels of DNA methyltransferase and DNA demethylase in gastric carcinomas." Oncol Rep **8**(5): 1085-9.
- Pal, S., S. N. Vishwanath, et al. (2004). "Human SWI/SNF-associated PRMT5 methylates histone H3 arginine 8 and negatively regulates expression of ST7 and NM23 tumor suppressor genes." Mol Cell Biol **24**(21): 9630-45.

- Palmer, B. R. and M. G. Marinus (1994). "The dam and dcm strains of *Escherichia coli*--a review." Gene **143**(1): 1-12.
- Pandey, M., S. Shukla, et al. "Promoter demethylation and chromatin remodeling by green tea polyphenols leads to re-expression of GSTP1 in human prostate cancer cells." Int J Cancer **126**(11): 2520-33.
- Peng, D. F., Y. Kanai, et al. (2005). "Increased DNA methyltransferase 1 (DNMT1) protein expression in precancerous conditions and ductal carcinomas of the pancreas." Cancer Sci **96**(7): 403-8.
- Peters, A. H., J. E. Mermoud, et al. (2002). "Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin." Nat Genet **30**(1): 77-80.
- Pheesse, T. J., L. Parry, et al. (2008). "Deficiency of Mbd2 attenuates Wnt signaling." Mol Cell Biol **28**(19): 6094-103.
- Pogo, B. G., V. G. Allfrey, et al. (1966). "RNA synthesis and histone acetylation during the course of gene activation in lymphocytes." Proc Natl Acad Sci U S A **55**(4): 805-12.
- Pogo, B. G., V. G. Allfrey, et al. (1967). "The effect of phytohemagglutinin on ribonucleic acid synthesis and histone acetylation in equine leukocytes." J Cell Biol **35**(2): 477-82.
- Pradhan, S., A. Bacolla, et al. (1999). "Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation." J Biol Chem **274**(46): 33002-10.
- Pradhan, S., D. Talbot, et al. (1997). "Baculovirus-mediated expression and characterization of the full-length murine DNA methyltransferase." Nucleic Acids Res **25**(22): 4666-73.
- Prokhortchouk, A., B. Hendrich, et al. (2001). "The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor." Genes Dev **15**(13): 1613-8.
- Prokhortchouk, A., O. Sansom, et al. (2006). "Kaiso-deficient mice show resistance to intestinal cancer." Mol Cell Biol **26**(1): 199-208.
- Ptashne, M. (2007). "On the use of the word 'epigenetic'." Curr Biol **17**(7): R233-6.
- Puig, O. M., E. Belles, et al. (1998). "Interaction between N-terminal domain of H4 and DNA is regulated by the acetylation degree." Biochim Biophys Acta **1397**(1): 79-90.
- Pulukuri, S. M., N. Estes, et al. (2007). "Demethylation-linked activation of urokinase plasminogen activator is involved in progression of prostate cancer." Cancer Res **67**(3): 930-9.
- Ramsahoye, B. H., D. Biniszkiewicz, et al. (2000). "Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a." Proc Natl Acad Sci U S A **97**(10): 5237-42.
- Rea, S., F. Eisenhaber, et al. (2000). "Regulation of chromatin structure by site-specific histone H3 methyltransferases." Nature **406**(6796): 593-9.

- Reinke, H. and W. Horz (2003). "Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter." Mol Cell **11**(6): 1599-607.
- Rhee, I., K. E. Bachman, et al. (2002). "DNMT1 and DNMT3b cooperate to silence genes in human cancer cells." Nature **416**(6880): 552-6.
- Rhee, I., K. W. Jair, et al. (2000). "CpG methylation is maintained in human cancer cells lacking DNMT1." Nature **404**(6781): 1003-7.
- Ricciardi, S., C. Kilstrup-Nielsen, et al. (2009). "CDKL5 influences RNA splicing activity by its association to the nuclear speckle molecular machinery." Hum Mol Genet **18**(23): 4590-602.
- Rice, J. C. and C. D. Allis (2001). "Histone methylation versus histone acetylation: new insights into epigenetic regulation." Curr Opin Cell Biol **13**(3): 263-73.
- Rice, J. C., S. D. Briggs, et al. (2003). "Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains." Mol Cell **12**(6): 1591-8.
- Robertson, K. D., E. Uzvolgyi, et al. (1999). "The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors." Nucleic Acids Res **27**(11): 2291-8.
- Rodova, M., K. F. Kelly, et al. (2004). "Regulation of the rapsyn promoter by kaiso and delta-catenin." Mol Cell Biol **24**(16): 7188-96.
- Rodriguez, J., L. Vives, et al. (2008). "Genome-wide tracking of unmethylated DNA Alu repeats in normal and cancer cells." Nucleic Acids Res **36**(3): 770-84.
- Rodriguez, P., E. Bonte, et al. (2005). "GATA-1 forms distinct activating and repressive complexes in erythroid cells." Embo J **24**(13): 2354-66.
- Rohrs, S., W. G. Dirks, et al. (2009). "Hypomethylation and expression of BEX2, IGSF4 and TIMP3 indicative of MLL translocations in acute myeloid leukemia." Mol Cancer **8**: 86.
- Ron-Bigger, S., O. Bar-Nur, et al. "Aberrant Epigenetic Silencing of Tumor Suppressor Genes is Reversed by Direct Reprogramming." Stem Cells.
- Roth, S. Y., J. M. Denu, et al. (2001). "Histone acetyltransferases." Annu Rev Biochem **70**: 81-120.
- Rottach, A., C. Frauer, et al. "The multi-domain protein Np95 connects DNA methylation and histone modification." Nucleic Acids Res **38**(6): 1796-804.
- Rougeulle, C., J. Chaumeil, et al. (2004). "Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome." Mol Cell Biol **24**(12): 5475-84.
- Ruas, M. and G. Peters (1998). "The p16INK4a/CDKN2A tumor suppressor and its relatives." Biochim Biophys Acta **1378**(2): F115-77.

- Rudolph-Owen, L. A., R. Chan, et al. (1998). "The matrix metalloproteinase matrilysin influences early-stage mammary tumorigenesis." Cancer Res **58**(23): 5500-6.
- Rush, M., R. Appanah, et al. (2009). "Targeting of EZH2 to a defined genomic site is sufficient for recruitment of Dnmt3a but not de novo DNA methylation." Epigenetics **4**(6): 404-14.
- Ruzov, A., D. S. Dunican, et al. (2004). "Kaiso is a genome-wide repressor of transcription that is essential for amphibian development." Development **131**(24): 6185-94.
- Ruzov, A., E. Savitskaya, et al. (2009). "The non-methylated DNA-binding function of Kaiso is not required in early *Xenopus laevis* development." Development **136**(5): 729-38.
- Saito, Y., Y. Kanai, et al. (2003). "Increased protein expression of DNA methyltransferase (DNMT) 1 is significantly correlated with the malignant potential and poor prognosis of human hepatocellular carcinomas." Int J Cancer **105**(4): 527-32.
- Saito, Y., Y. Kanai, et al. (2002). "Overexpression of a splice variant of DNA methyltransferase 3b, DNMT3b4, associated with DNA hypomethylation on pericentromeric satellite regions during human hepatocarcinogenesis." Proc Natl Acad Sci U S A **99**(15): 10060-5.
- Sankar, N., S. Baluchamy, et al. (2008). "p300 provides a corepressor function by cooperating with YY1 and HDAC3 to repress c-Myc." Oncogene **27**(43): 5717-28.
- Sansom, O. J., J. Berger, et al. (2003). "Deficiency of Mbd2 suppresses intestinal tumorigenesis." Nat Genet **34**(2): 145-7.
- Sansom, O. J., J. Zabkiewicz, et al. (2003). "MBD4 deficiency reduces the apoptotic response to DNA-damaging agents in the murine small intestine." Oncogene **22**(46): 7130-6.
- Santos-Rosa, H., R. Schneider, et al. (2002). "Active genes are tri-methylated at K4 of histone H3." Nature **419**(6905): 407-11.
- Santos, F., B. Hendrich, et al. (2002). "Dynamic reprogramming of DNA methylation in the early mouse embryo." Dev Biol **241**(1): 172-82.
- Sarraf, S. A. and I. Stancheva (2004). "Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly." Mol Cell **15**(4): 595-605.
- Sasai, N. and P. A. Defossez (2009). "Many paths to one goal? The proteins that recognize methylated DNA in eukaryotes." Int J Dev Biol **53**(2-3): 323-34.
- Sasai, N., E. Matsuda, et al. (2005). "Identification of a novel BTB-zinc finger transcriptional repressor, CIBZ, that interacts with CtBP corepressor." Genes Cells **10**(9): 871-85.
- Sasai, N., M. Nakao, et al. "Sequence-specific recognition of methylated DNA by human zinc-finger proteins." Nucleic Acids Res.

- Sato, M., Y. Horio, et al. (2002). "The expression of DNA methyltransferases and methyl-CpG-binding proteins is not associated with the methylation status of p14(ARF), p16(INK4a) and RASSF1A in human lung cancer cell lines." Oncogene **21**(31): 4822-9.
- Sawada, M., Y. Kanai, et al. (2007). "Increased expression of DNA methyltransferase 1 (DNMT1) protein in uterine cervix squamous cell carcinoma and its precursor lesion." Cancer Lett **251**(2): 211-9.
- Saxonov, S., P. Berg, et al. (2006). "A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters." Proc Natl Acad Sci U S A **103**(5): 1412-7.
- Schmidt, W. M., R. Sedivy, et al. (2007). "Progressive up-regulation of genes encoding DNA methyltransferases in the colorectal adenoma-carcinoma sequence." Mol Carcinog **46**(9): 766-72.
- Schneider, R., A. J. Bannister, et al. (2004). "Histone H3 lysine 4 methylation patterns in higher eukaryotic genes." Nat Cell Biol **6**(1): 73-7.
- Schones, D. E., K. Cui, et al. (2008). "Dynamic regulation of nucleosome positioning in the human genome." Cell **132**(5): 887-98.
- Schotta, G., M. Lachner, et al. (2004). "A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin." Genes Dev **18**(11): 1251-62.
- Schule, B., H. H. Li, et al. (2007). "DLX5 and DLX6 expression is biallelic and not modulated by MeCP2 deficiency." Am J Hum Genet **81**(3): 492-506.
- Schulz, W. A., J. P. Elo, et al. (2002). "Genomewide DNA hypomethylation is associated with alterations on chromosome 8 in prostate carcinoma." Genes Chromosomes Cancer **35**(1): 58-65.
- Screaton, R. A., S. Kiessling, et al. (2003). "Fas-associated death domain protein interacts with methyl-CpG binding domain protein 4: a potential link between genome surveillance and apoptosis." Proc Natl Acad Sci U S A **100**(9): 5211-6.
- Sealy, L. and R. Chalkley (1978). "DNA associated with hyperacetylated histone is preferentially digested by DNase I." Nucleic Acids Res **5**(6): 1863-76.
- Sekimata, M. and Y. Homma (2004). "Regulation of Rb gene expression by an MBD2-interacting zinc finger protein MIZF during myogenic differentiation." Biochem Biophys Res Commun **325**(3): 653-9.
- Sekimata, M. and Y. Homma (2004). "Sequence-specific transcriptional repression by an MBD2-interacting zinc finger protein MIZF." Nucleic Acids Res **32**(2): 590-7.
- Sekimata, M., A. Takahashi, et al. (2001). "Involvement of a novel zinc finger protein, MIZF, in transcriptional repression by interacting with a methyl-CpG-binding protein, MBD2." J Biol Chem **276**(46): 42632-8.

- Sekiyama, N., T. Ikegami, et al. (2008). "Structure of the small ubiquitin-like modifier (SUMO)-interacting motif of MBD1-containing chromatin-associated factor 1 bound to SUMO-3." J Biol Chem **283**(51): 35966-75.
- Selker, E. U. (1998). "Trichostatin A causes selective loss of DNA methylation in Neurospora." Proc Natl Acad Sci U S A **95**(16): 9430-5.
- Selker, E. U., N. A. Tountas, et al. (2003). "The methylated component of the Neurospora crassa genome." Nature **422**(6934): 893-7.
- Serre, D., B. H. Lee, et al. "MBD-isolated Genome Sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome." Nucleic Acids Res **38**(2): 391-9.
- Shann, Y. J., C. Cheng, et al. (2008). "Genome-wide mapping and characterization of hypomethylated sites in human tissues and breast cancer cell lines." Genome Res **18**(5): 791-801.
- Sharif, J., M. Muto, et al. (2007). "The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA." Nature **450**(7171): 908-12.
- Shi, Y., F. Lan, et al. (2004). "Histone demethylation mediated by the nuclear amine oxidase homolog LSD1." Cell **119**(7): 941-53.
- Shi, Y., J. Sawada, et al. (2003). "Coordinated histone modifications mediated by a CtBP co-repressor complex." Nature **422**(6933): 735-8.
- Shukeir, N., P. Pakneshan, et al. (2006). "Alteration of the methylation status of tumor-promoting genes decreases prostate cancer cell invasiveness and tumorigenesis in vitro and in vivo." Cancer Res **66**(18): 9202-10.
- Skene, P. J., R. S. Illingworth, et al. "Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state." Mol Cell **37**(4): 457-68.
- Slack, A., V. Bovenzi, et al. (2002). "Antisense MBD2 gene therapy inhibits tumorigenesis." J Gene Med **4**(4): 381-9.
- Smiraglia, D. J., L. J. Rush, et al. (2001). "Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies." Hum Mol Genet **10**(13): 1413-9.
- Soares, J., A. E. Pinto, et al. (1999). "Global DNA hypomethylation in breast carcinoma: correlation with prognostic factors and tumor progression." Cancer **85**(1): 112-8.
- Song, J. Z., C. Stirzaker, et al. (2002). "Hypermethylation trigger of the glutathione-S-transferase gene (GSTP1) in prostate cancer cells." Oncogene **21**(7): 1048-61.
- Soubry, A., J. van Hengel, et al. (2005). "Expression and nuclear location of the transcriptional repressor Kaiso is regulated by the tumor microenvironment." Cancer Res **65**(6): 2224-33.

- Spada, F., A. Haemmer, et al. (2007). "DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells." J Cell Biol **176**(5): 565-71.
- Spring, C. M., K. F. Kelly, et al. (2005). "The catenin p120ctn inhibits Kaiso-mediated transcriptional repression of the beta-catenin/TCF target gene matrilysin." Exp Cell Res **305**(2): 253-65.
- Spruijt, C. G., S. J. Bartels, et al. "CDK2AP1/DOC-1 is a bona fide subunit of the Mi-2/NuRD complex." Mol Biosyst.
- Squillaro, T., N. Alessio, et al. "Partial silencing of methyl cytosine protein binding 2 (MECP2) in mesenchymal stem cells induces senescence with an increase in damaged DNA." Faseb J **24**(5): 1593-603.
- Steger, D. J., M. I. Lefterova, et al. (2008). "DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells." Mol Cell Biol **28**(8): 2825-39.
- Sterner, D. E. and S. L. Berger (2000). "Acetylation of histones and transcription-related factors." Microbiol Mol Biol Rev **64**(2): 435-59.
- Stirzaker, C., J. Z. Song, et al. (2004). "Transcriptional gene silencing promotes DNA hypermethylation through a sequential change in chromatin modifications in cancer cells." Cancer Res **64**(11): 3871-7.
- Strahl, B. D. and C. D. Allis (2000). "The language of covalent histone modifications." Nature **403**(6765): 41-5.
- Suetake, I., F. Shinozaki, et al. (2004). "DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction." J Biol Chem **279**(26): 27816-23.
- Surani, M. A. (1998). "Imprinting and the initiation of gene silencing in the germ line." Cell **93**(3): 309-12.
- Suzuki, M., N. Sunaga, et al. (2004). "RNA interference-mediated knockdown of DNA methyltransferase 1 leads to promoter demethylation and gene re-expression in human lung and breast cancer cells." Cancer Res **64**(9): 3137-43.
- Suzuki, M., T. Yamada, et al. (2003). "Direct association between PU.1 and MeCP2 that recruits mSin3A-HDAC complex for PU.1-mediated transcriptional repression." Oncogene **22**(54): 8688-98.
- Swanberg, S. E., R. P. Nagarajan, et al. (2009). "Reciprocal co-regulation of EGR2 and MECP2 is disrupted in Rett syndrome and autism." Hum Mol Genet **18**(3): 525-34.
- Tahiliani, M., P. Mei, et al. (2007). "The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation." Nature **447**(7144): 601-5.
- Takai, D. and P. A. Jones (2002). "Comprehensive analysis of CpG islands in human chromosomes 21 and 22." Proc Natl Acad Sci U S A **99**(6): 3740-5.

- Takeshima, H., S. Yamashita, et al. (2009). "The presence of RNA polymerase II, active or stalled, predicts epigenetic fate of promoter CpG islands." Genome Res **19**(11): 1974-82.
- Tan, C. P. and S. Nakielny (2006). "Control of the DNA methylation system component MBD2 by protein arginine methylation." Mol Cell Biol **26**(19): 7224-35.
- Tao, J., K. Hu, et al. (2009). "Phosphorylation of MeCP2 at Serine 80 regulates its chromatin association and neurological function." Proc Natl Acad Sci U S A **106**(12): 4882-7.
- Tatematsu, K. I., T. Yamazaki, et al. (2000). "MBD2-MBD3 complex binds to hemimethylated DNA and forms a complex containing DNMT1 at the replication foci in late S phase." Genes Cells **5**(8): 677-88.
- Taunton, J., C. A. Hassig, et al. (1996). "A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p." Science **272**(5260): 408-11.
- Thomas, J. O. and R. D. Kornberg (1975). "An octamer of histones in chromatin and free in solution." Proc Natl Acad Sci U S A **72**(7): 2626-30.
- Thomson, J. P., P. J. Skene, et al. "CpG islands influence chromatin structure via the CpG-binding protein Cfp1." Nature **464**(7291): 1082-6.
- Tian, F., X. Z. Hu, et al. (2009). "Dynamic chromatin remodeling events in hippocampal neurons are associated with NMDA receptor-mediated activation of Bdnf gene promoter 1." J Neurochem **109**(5): 1375-88.
- Tillo, D. and T. R. Hughes (2009). "G+C content dominates intrinsic nucleosome occupancy." BMC Bioinformatics **10**: 442.
- Ting, A. H., K. W. Jair, et al. (2004). "CpG island hypermethylation is maintained in human colorectal cancer cells after RNAi-mediated depletion of DNMT1." Nat Genet **36**(6): 582-4.
- Tiwari, V. K., K. M. McGarvey, et al. (2008). "PcG proteins, DNA methylation, and gene repression by chromatin looping." PLoS Biol **6**(12): 2911-27.
- Trasler, J. M., D. G. Trasler, et al. (1996). "DNA methyltransferase in normal and Dnmtn/Dnmtn mouse embryos." Dev Dyn **206**(3): 239-47.
- Trojer, P., G. Li, et al. (2007). "L3MBTL1, a histone-methylation-dependent chromatin lock." Cell **129**(5): 915-28.
- Tryndyak, V. P., O. Kovalchuk, et al. (2006). "Loss of DNA methylation and histone H4 lysine 20 trimethylation in human breast cancer cells is associated with aberrant expression of DNA methyltransferase 1, Suv4-20h2 histone methyltransferase and methyl-binding proteins." Cancer Biol Ther **5**(1): 65-70.
- Tsujimura, K., M. Abematsu, et al. (2009). "Neuronal differentiation of neural precursor cells is promoted by the methyl-CpG-binding protein MeCP2." Exp Neurol **219**(1): 104-11.

- Tsukada, Y., J. Fang, et al. (2006). "Histone demethylation by a family of JmjC domain-containing proteins." Nature **439**(7078): 811-6.
- Tudor, M., S. Akbarian, et al. (2002). "Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain." Proc Natl Acad Sci U S A **99**(24): 15536-41.
- Turner, B. M. (1993). "Decoding the nucleosome." Cell **75**(1): 5-8.
- Uchimura, Y., T. Ichimura, et al. (2006). "Involvement of SUMO modification in MBD1- and MCAF1-mediated heterochromatin formation." J Biol Chem **281**(32): 23180-90.
- Unoki, M., T. Nishidate, et al. (2004). "ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain." Oncogene **23**(46): 7601-10.
- Urduingio, R. G., L. Lopez-Serra, et al. (2008). "Mecp2-null mice provide new neuronal targets for Rett syndrome." PLoS One **3**(11): e3669.
- Vecsler, M., A. J. Simon, et al. "MeCP2 deficiency downregulates specific nuclear proteins that could be partially recovered by valproic acid in vitro." Epigenetics **5**(1): 61-7.
- Vempati, R. K., R. S. Jayani, et al. "p300 mediated acetylation of histone H3 lysine 56 functions in DNA damage response in mammals." J Biol Chem.
- Vertino, P. M., R. W. Yen, et al. (1996). "De novo methylation of CpG island sequences in human fibroblasts overexpressing DNA (cytosine-5-)-methyltransferase." Mol Cell Biol **16**(8): 4555-65.
- Villa, R., L. Morey, et al. (2006). "The methyl-CpG binding protein MBD1 is required for PML-RARalpha function." Proc Natl Acad Sci U S A **103**(5): 1400-5.
- Visel, A., M. J. Blow, et al. (2009). "ChIP-seq accurately predicts tissue-specific activity of enhancers." Nature **457**(7231): 854-8.
- Vogel, M. J., D. Peric-Hupkes, et al. (2007). "Detection of in vivo protein-DNA interactions using DamID in mammalian cells." Nat Protoc **2**(6): 1467-78.
- Wade, P. A., D. Pruss, et al. (1997). "Histone acetylation: chromatin in action." Trends Biochem Sci **22**(4): 128-32.
- Wang, C., J. Xu, et al. "An effective approach for identification of in vivo protein-DNA binding sites from paired-end ChIP-Seq data." BMC Bioinformatics **11**: 81.
- Wang, H., Z. Q. Huang, et al. (2001). "Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor." Science **293**(5531): 853-7.
- Wang, Z., C. Zang, et al. (2009). "Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes." Cell **138**(5): 1019-31.

- Wang, Z., C. Zang, et al. (2008). "Combinatorial patterns of histone acetylations and methylations in the human genome." Nat Genet **40**(7): 897-903.
- Watanabe, S., T. Ichimura, et al. (2003). "Methylated DNA-binding domain 1 and methylpurine-DNA glycosylase link transcriptional repression and DNA repair in chromatin." Proc Natl Acad Sci U S A **100**(22): 12859-64.
- Weber, A., J. Marquardt, et al. (2008). "Zbtb4 represses transcription of P21CIP1 and controls the cellular response to p53 activation." Embo J **27**(11): 1563-74.
- Weber, M., J. J. Davies, et al. (2005). "Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells." Nat Genet **37**(8): 853-62.
- Weber, M., I. Hellmann, et al. (2007). "Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome." Nat Genet **39**(4): 457-66.
- Weidemann, T., M. Wachsmuth, et al. (2003). "Counting nucleosomes in living cells with a combination of fluorescence correlation spectroscopy and confocal imaging." J Mol Biol **334**(2): 229-40.
- Weller, M., R. Stupp, et al. "MGMT promoter methylation in malignant gliomas: ready for personalized medicine?" Nat Rev Neurol **6**(1): 39-51.
- Wen, X. Z., Y. Akiyama, et al. "Methylation of GATA-4 and GATA-5 and development of sporadic gastric carcinomas." World J Gastroenterol **16**(10): 1201-8.
- Whetstine, J. R., A. Nottke, et al. (2006). "Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases." Cell **125**(3): 467-81.
- Wild, L., J. M. Funes, et al. "In Vitro Transformation of Mesenchymal Stem Cells Induces Gradual Genomic Hypomethylation." Carcinogenesis.
- Winter, S., W. Fischle, et al. (2008). "Modulation of 14-3-3 interaction with phosphorylated histone H3 by combinatorial modification patterns." Cell Cycle **7**(10): 1336-42.
- Winter, S., E. Simboeck, et al. (2008). "14-3-3 proteins recognize a histone code at histone H3 and are required for transcriptional activation." Embo J **27**(1): 88-99.
- Wolffe, A. P. and J. J. Hayes (1999). "Chromatin disruption and modification." Nucleic Acids Res **27**(3): 711-20.
- Wong, E., K. Yang, et al. (2002). "Mbd4 inactivation increases Cright-arrowT transition mutations and promotes gastrointestinal tumor formation." Proc Natl Acad Sci U S A **99**(23): 14937-42.
- Woo, H. R., T. A. Dittmer, et al. (2008). "Three SRA-domain methylcytosine-binding proteins cooperate to maintain global CpG methylation and epigenetic silencing in Arabidopsis." PLoS Genet **4**(8): e1000156.
- Wood, L. and G. M. Shepherd "Synaptic circuit abnormalities of motor-frontal layer 2/3 pyramidal neurons in a mutant mouse model of Rett syndrome." Neurobiol Dis **38**(2): 281-7.

- Wossidlo, M., J. Arand, et al. "Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes." *Embo J* **29**(11): 1877-88.
- Wu, B. and C. A. Davey "Using Soft X-Rays for a Detailed Picture of Divalent Metal Binding in the Nucleosome." *J Mol Biol.*
- Xie, X., T. Kokubo, et al. (1996). "Structural similarity between TAFs and the heterotetrameric core of the histone octamer." *Nature* **380**(6572): 316-22.
- Xu, G. L., T. H. Bestor, et al. (1999). "Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene." *Nature* **402**(6758): 187-91.
- Yamane, K., C. Toumazou, et al. (2006). "JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor." *Cell* **125**(3): 483-95.
- Yang, X. J. and E. Seto (2008). "The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men." *Nat Rev Mol Cell Biol* **9**(3): 206-18.
- Yasui, D. H., S. Peddada, et al. (2007). "Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes." *Proc Natl Acad Sci U S A* **104**(49): 19416-21.
- Yegnasubramanian, S., X. Lin, et al. (2006). "Combination of methylated-DNA precipitation and methylation-sensitive restriction enzymes (COMPARE-MS) for the rapid, sensitive and quantitative detection of DNA methylation." *Nucleic Acids Res* **34**(3): e19.
- Yen, R. W., P. M. Vertino, et al. (1992). "Isolation and characterization of the cDNA encoding human DNA methyltransferase." *Nucleic Acids Res* **20**(9): 2287-91.
- Yoder, J. A., R. W. Yen, et al. (1996). "New 5' regions of the murine and human genes for DNA (cytosine-5)-methyltransferase." *J Biol Chem* **271**(49): 31092-7.
- Yoon, H. G., D. W. Chan, et al. (2003). "N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso." *Mol Cell* **12**(3): 723-34.
- Yoshiura, K., Y. Kanai, et al. (1995). "Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas." *Proc Natl Acad Sci U S A* **92**(16): 7416-9.
- Young, J. I., E. P. Hong, et al. (2005). "Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2." *Proc Natl Acad Sci U S A* **102**(49): 17551-8.
- Yu, F., J. Thiesen, et al. (2000). "Histone deacetylase-independent transcriptional repression by methyl-CpG-binding protein 2." *Nucleic Acids Res* **28**(10): 2201-6.
- Yuan, J., M. Pu, et al. (2009). "Histone H3-K56 acetylation is important for genomic stability in mammals." *Cell Cycle* **8**(11): 1747-53.

Zhang, X., Y. Ozawa, et al. (2005). "Histone deacetylase 3 (HDAC3) activity is regulated by interaction with protein serine/threonine phosphatase 4." Genes Dev **19**(7): 827-39.

Zhang, Y., H. H. Ng, et al. (1999). "Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation." Genes Dev **13**(15): 1924-35.

Zhao, X., T. Ueba, et al. (2003). "Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function." Proc Natl Acad Sci U S A **100**(11): 6777-82.

Zhou, Z., E. J. Hong, et al. (2006). "Brain-specific phosphorylation of MeCP2 regulates activity-dependent Bdnf transcription, dendritic growth, and spine maturation." Neuron **52**(2): 255-69.

Annexes

Annexe 1

Specific association between the methyl-CpG-binding domain protein 2 and the hypermethylated region of the human telomerase reverse transcriptase promoter in cancer cells

Amandine Chatagnon^{1,†}, Stéphanie Bougel^{2,†}, Laury Perriaud¹, Joël Lachuer³, Jean Benhattar² and Robert Dante^{1,*}

¹INSERM U590, Lyon F-69008, France, ²Institut de Pathologie, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Lausanne CH-1011, Switzerland and ³ProfilExpert/Neurobiotec Service, INSERM U842, Bron F-69676, France

*To whom correspondence should be addressed. INSERM U590, Oncogénèse et Progression Tumorale, Centre Léon Bérard, 28 rue Laënnec, 69373 Lyon Cedex 08, France. Tel: +33 4 78 78 59 22; Fax: +33 4 78 78 27 20; Email: dante@univ-lyon1.fr

Human telomerase reverse transcriptase (*hTERT*) is expressed in most cancer cells. Paradoxically, its promoter is embedded in a hypermethylated CpG island. A short region escapes to this alteration, allowing a basal level of transcription. However, the methylation of adjacent regions may play a role in the maintenance of low *hTERT* expression. It is now well established that methyl-CpG binding domain proteins mediate the transcriptional silencing of hypermethylated genes. The potential involvement of these proteins in the control of *hTERT* expression was firstly investigated in HeLa cells. Chromatin immunoprecipitation assays showed that only methyl-CpG-binding domain protein 2 (MBD2) associated the hypermethylated *hTERT* promoter. In MBD2 knock-down HeLa cells, constitutively depleted in MBD2, neither methyl CpG binding protein 2 (MeCP2) nor MBD1 acted as substitutes for MBD2. MBD2 depletion by transient or constitutive RNA interference led to an upregulation of *hTERT* transcription that can be downregulated by expressing mouse Mbd2 protein. Our results indicate that MBD2 is specifically and directly involved in the transcriptional repression of *hTERT* in HeLa cells. This specific transcriptional repression was also observed in breast, liver and neuroblastoma cancer cell lines. Thus, MBD2 seems to be a general repressor of *hTERT* in *hTERT*-methylated telomerase-positive cells.

Introduction

An increasing body of evidences indicate that the alterations of DNA methylation patterns are a characteristic of cancer cells (1). Generally, global reduction of DNA methylation level is associated with localized hypermethylation (1). Specifically, an aberrant hypermethylation of CpG islands at the 5' end of tumor suppressor genes, leading to transcriptional repression, has been described both in cancer cell lines and tumor tissues (2,3).

In human epithelial and fibroblast cells, telomere shortening is a key event in replicative senescence. In >85% of cancer cells, telomere length is maintained through telomerase holoenzyme activity (4,5). Although germ cells and stem cells also exhibit high telomerase activity (4), in normal somatic cells, the catalytic subunit of the telomerase [*human telomerase reverse transcriptase (hTERT)*] is silenced, leading to a limited life span (6). The regulation of *hTERT* transcription has been extensively investigated and several inducers and repressors have been identified, including c-Myc, Sp1, hALP, Hif-1, Mbi-1, USF1/2, estrogen response element, p53, Mad1, myeloid-

Abbreviations: ChIP, chromatin immunoprecipitation; hTERT, human telomerase reverse transcriptase; KD, knockdown; MBD, methyl-CpG-binding domain; mRNA, messenger RNA; nt, nucleotide; PCR, polymerase chain reaction; RT, reverse transcription.

[†]These authors contributed equally to this work.

specific zinc finger protein 2, transforming growth factor- β , Wilms' Tumor 1 and CTCF (7–9). In addition, the *hTERT* promoter region is embedded in a large CpG island spanning nucleotides (nt) –1100 to +1500 from the transcription start site, suggesting that transcription of the gene might be regulated by DNA methylation.

The first studies examining the methylation status of the *hTERT* CpG island had led to a paradox. In normal somatic cells, this CpG island was found unmethylated while the gene was transcriptionally silent. However, in most of cancer cells, this region was aberrantly methylated, whereas telomerase activities and *hTERT* messenger RNAs (mRNAs) were unambiguously detected (10–13). This paradox was recently solved. *hTERT* methylation prevents the binding of negatively acting transcription factors such as CTCF inhibitor (9), and a partial hypomethylation of the *hTERT* promoter region can result in some level of transcriptional activity (14). In several cancer cell lines and tumor tissues, careful analysis of *hTERT* methylation patterns has shown that a short region of the CpG island (positions nt –165 to nt –80) is unmethylated or slightly methylated despite highly methylated border regions (14). This unmethylated region is located in the *hTERT* core promoter (positions nt –279 to nt +5) (15), and chromatin immunoprecipitation (ChIP) assays have shown that active chromatin marks are associated with this unmethylated region (16). Moreover, studies using plasmid expression vectors and patch methylation techniques indicate that the *hTERT* core promoter does not show any activity when all CpG sites are methylated (17). In contrast, the selective demethylation of a small region upstream the transcription start site significantly activates the *hTERT* promoter in a reporter plasmid. Nevertheless, the activity of the promoter under these conditions is significantly lower than when using the unmethylated core promoter. In this condition, *hTERT* does not provide a real exception to the general model of gene silencing by promoter methylation and the hypermethylation around the unmethylated region seems to play a major role in the reduction of *hTERT* transcriptional activity. Indeed, in telomerase-positive cancer cells, only 0.2–6 mRNA molecules per cell can be detected (18,19), suggesting that the transcriptional activity of the *hTERT* promoter is limited by cellular factors since high transcription rates are induced when the unmethylated promoter is used in plasmid expression vectors (14).

Transcriptional repression mediated by CpG methylation often involves methyl-CpG-binding domain (MBD) proteins. The five MBD proteins identified to date, methyl CpG binding protein 2 (MeCP2), MBD1, MBD2, MBD3 and MBD4, share a highly conserved MBD. With the exception of MBD4, which is primarily a thymine glycosylase involved in DNA repair (20), all MBD proteins are involved in the transcriptional repression mediated by DNA methylation. It has now been well established that MeCP2, MBD1 and MBD2 bind to methylated DNA and recruit different histone deacetylase complexes and histone methyl transferases, belonging to the chromatin remodeling complexes that control chromatin compaction and induce gene silencing (21–23). MBD3 lacks a functional MBD but is an integral subunit of histone deacetylase complex the Mi2–NuRD complex that is recruited through MBD2 (24,25).

These data have prompted us to investigate whether MBD proteins are involved in the repression of *hTERT* expression when hypermethylated in telomerase-positive cells.

Materials and methods

Cell culture

Five human tumor cell lines, HeLa (cervical adenocarcinoma), MCF7 (breast adenocarcinoma), HepG2 (hepatocellular carcinoma), LAN-1 (neuroblastoma) and NCCIT (teratocarcinoma) and a normal human embryonic lung fibroblast

cell line, MRC5, were used in this study. Cells were obtained from the American Type Culture Collection (Rockville, MD), except for LAN-1, and grown in the medium recommended at 37°C in a humidified 5% CO₂ atmosphere.

ChIP assays

ChIP was done as described previously (26). Cross-linked chromatin was immunoprecipitated using 15 µl of two different polyclonal anti-MBD2 antibodies (kindly provided by Dr P.Wade and Dr E.Ballestar) or 20 µl of polyclonal anti-MeCP2 (Upstate Biotechnology, Lake Placid, NY), anti-MBD1 (Abcam, Paris, France) antibodies or anti-mouse IgG (Dakocytomation, Trappes, France).

Purified DNAs obtained from the input, unbound and bound fractions were quantified by densitometry using the VersaFluor™ Fluorometer (Bio-rad, Ivry, France) and RiboGreen reagent (Molecular Probes, Interchim, Montluçon, France).

Polymerase chain reaction (PCR) analysis was then performed to assess the binding of MBD proteins to the *hTERT* core promoter. An equal quantity (0.4 ng) of each DNA fraction was amplified by dose-dependent PCR (supplementary Figure S1 is available at *Carcinogenesis* Online) using HotStar Taq polymerase Kit (Qiagen, Courtabouef, France), 5% of dimethyl sulfoxide and 0.5 µM of primers spanning a region from nt -296 to nt -84 of *hTERT* (*hTERT* ChIP, see supplementary Table S1 available at *Carcinogenesis* Online). The thermal cycler program was 37 cycles of 94°C for 30 s, 65°C for 60 s and 72°C for 90s. PCR products were analyzed on a 2% agarose gel containing 1 µg/ml ethidium bromide and were quantified by densitometry using a Fluor's fluorimeter and Quantity One software (Bio-rad).

ChIP-on-chip

For ChIP-on-chip analysis, the specific protein-DNA complexes were obtained from independent immunoprecipitations using two different polyclonal anti-MBD2 antibodies (kindly provided by Dr P.Wade and Dr E.Ballestar). The ChIP DNAs from the input and bound fractions were amplified, labeled and hybridized on microarrays by ProfileXpert service according to Affymetrix™ protocols. Briefly, the ChIP DNA was amplified by ligation-mediated PCR. To test for enrichment of MBD2-bound sites, PCR amplification of *NBR2* (26) and *pS2* promoters was performed on each ChIP samples before and after amplification. The amplified DNAs were then labeled using the GeneChip® WT Double-Stranded DNA Terminal Labelling Kit and hybridized to the human tiling arrays (Human Promoter 1.0R Arrays), which were then washed and scanned. Raw data from the scans were analyzed using Affymetrix® Tiling Analysis Software and the results were viewed in Affymetrix® Integrated Genome Browser Software.

DNA methylation analysis

Bisulfite sequencing used to determine the CpG methylation pattern of *hTERT* promoter and proximal exonic region was performed as described previously (12,14). Briefly, bisulfite-modified genomic DNA was amplified by two primer sets [P1, positions nt -442 to nt -219 (12) and P2, positions nt -206 to nt +108 (14)] to analyze a region from nt -442 to nt +108 from the transcription start site of *hTERT*. PCR amplifications were accomplished using the master mix (Promega, Madison, WI), with the following conditions: 40 cycles of 94°C for 30 s, 54°C (P1) or 57°C (P2) for 45 s and 72°C for 50 s. DNA methylation status was then established by a direct sequencing of PCR products or by a sequencing of cloned PCR products. For the last procedure, PCR products were cloned into the pGEM-T vector using the pGEM-T vector system II (Promega). After transformation of JM109 *Escherichia coli* competent cells (Promega), plasmid DNA was extracted from clones with the QIAprep Spin Miniprep Kit (Qiagen). Each clone was sequenced with the M13 forward primer (5'-GTAAACGACGGCCAG-3'), using a Big Dye Terminator Cycle Sequencing Kit and an ABI Prism 3100 sequencer (Applied Biosystems, Foster City, CA).

Transient transfection

siRNA duplexes for *MBD2* (sense: 5'-GGAGGAAGUGUACCGAAAATT-3' and antisense: 5'-UUUUCGGAUCACUCCUCCTT-3'; Eurogentec, Seraing, Belgium); non-specific small interfering RNA (siRNA) control (Eurogentec); pRev-Mbd2 (kindly provided by Dr A.Bird) (26) and pGL3 basic (Promega) were transfected in cell lines using lipofectamine 2000 (Invitrogen, Carlsbad, CA) according to the manufacturer's recommendations. Briefly, cells were seeded at 2 × 10⁵ cells per well in six-well plates and grown to 50–60% confluence on the day of transfection. All transfections were done in Opti-MEM medium (Invitrogen) with 625 nM of *MBD2* siRNA or 1 µg of *Mbd2* expression plasmid. Lipofectamine 2000 complexes were incubated for 4–5 h. The medium was then removed and replaced with fresh medium. Cells were grown and harvested at various time after the transfection.

RNA extraction and reverse transcription-PCR analysis

Total RNA was extracted from cells using the RNeasy Mini Kit (Qiagen). After extraction, the integrity of total RNA was examined on a 1.2% agarose gel

containing 1 µg/ml ethidium bromide and quantified by densitometry using a Fluor's fluorimeter and Quantity One software (Bio-rad) by comparison with serial dilutions of a standard RNA (Roche Molecular Biochemicals, Maylan, France).

MBD2 mRNA was quantified by competitive quantitative reverse transcription (RT)-PCR as described previously (27). *hTERT* mRNA levels were monitored by relative RT-PCR using One Step RT-PCR kit (Qiagen) and 0.1 µg of total RNA. Cycling parameters were 50°C for 30 min followed by 95°C for 15 min and then 32 cycles of 94°C for 30 s, 55°C for 60 s and 72°C for 90 s. *hTERT* transcripts were co-amplified with *PBGD* transcripts used as internal controls (for primers *hTERT* RT-PCR and *PBGD* RT-PCR, see supplementary Table S1 available at *Carcinogenesis* Online). PCR products were analyzed on a 2% agarose gel containing 1 µg/ml ethidium bromide and quantified by densitometry. The ratio between *hTERT* and *PBGD* signals was determined. To quantify more precisely *hTERT* mRNA, real-time RT-PCR was also carried out on a Rotorgene 6000 cycler (Corbett Research, Sydney, Australia) using *C. therm.* Polymerase One-Step RT-PCR System (Roche Molecular Biochemicals, Maylan, France). Each reaction mixture included 50 ng of template RNA, 250 nM of FAM-labeled probe (see supplementary Table S1 available at *Carcinogenesis* Online) and 500 nM of primers specific for *hTERT* (*hTERT* RT-PCR Q, see supplementary Table S1 available at *Carcinogenesis* Online). Cycling parameters were 60°C for 30 min followed by 95°C for 5 min and then 45 cycles at 95°C for 15 s and 60°C for 60 s. β -Actin mRNA levels were used to normalize *hTERT* expression. The relative level of each mRNA was calculated on the basis of two standard curves using the relative quantification method.

Results

hTERT hypermethylated CpG island is selectively associated with *MBD2* in HeLa cells

To explore the potential involvement of MBD proteins in *hTERT* regulation, we have chosen cervical cancer cell line HeLa, as a first model. In these telomerase-positive tumor cells, as was shown in our earlier studies, the *hTERT* core promoter is also regionally hypermethylated (17) and might be a target of MBD proteins.

To address this issue, ChIP assays were performed using antibodies directed against MeCP2, MBD1 and MBD2. As a control, the fractions immunoprecipitated with a non-MBD protein-specific antibody (anti-mouse IgG) were also analyzed. The relative amounts of *hTERT* core promoter were determined by a dose-dependent PCR assay (supplementary Figure S1 available at *Carcinogenesis* Online) using a constant amount of DNA from each fractionation process: input, unbound and bound fractions. Representative data from at least three independent experiments are shown in Figure 1A. A strong enrichment in *hTERT* promoter was observed in the fraction immunoprecipitated by anti-MBD2 antibodies when compared with input or non-retained fractions (Figure 1A 'input', 'unbound' and 'IgG'). Since dose-dependent PCR assays were performed with a constant amount of DNA, these data strongly suggest that MBD2 is associated with the methylated region of the *hTERT* promoter. In contrast, analysis of the fractions immunoprecipitated by anti-MeCP2 or anti-MBD1 antibodies showed that these fractions were depleted in *hTERT* DNA (Figure 1A), indicating that these two proteins are not bound to *hTERT* promoter in HeLa cells. Furthermore, the depletion in *hTERT* DNA observed in the fractions bound by anti-MeCP2 and anti-MBD1 antibodies suggests that these two proteins are probably linked to other chromatin domains in HeLa cells.

Taken together, these data strongly suggest that the methylated regions adjacent to the unmethylated region of the *hTERT* core promoter are selectively associated with MBD2.

Recently, we have developed high-throughput analyses of MBD2-binding pattern using a ChIP-on-chip approach (A. Chatagnon, L. Perriaud, J. Lachuer and R. Dante, in preparation). DNAs obtained from the chromatin immunoprecipitated by anti-MBD2 antibodies were hybridized to Affymetrix Human Promoter 1.0R Array. On this chip, 25 500 human promoter regions tiled at 35 bp resolution are representing. Each promoter region covers ~7.5 kb upstream through 2.45 kb downstream of 5' transcription start site and for 1300 cancer-associated genes, additional 2.45 kb are also represented. Data obtained from two independent experiments performed with two different anti-MBD2 antibodies indicated that MBD2 not only bound the

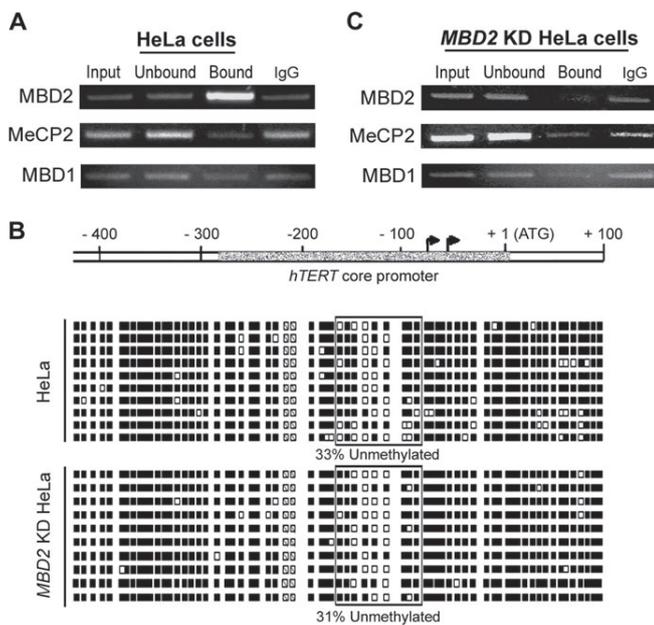


Fig. 1. MBD2 specifically binds the hypermethylated region of the *hTERT* promoter in HeLa cells. (A) ChIP analysis of MBD proteins binding to the hypermethylated region of the *hTERT* promoter (positions nt –296 to nt –84 from the ATG translational start site) in HeLa cells. Cross-linked chromatin was immunoprecipitated using anti-MeCP2, anti-MBD1 and anti-MBD2 antibodies or anti-mouse IgG. The relative amounts of *hTERT* core promoter were determined by a dose-dependent PCR assay using a constant amount of DNA from each step of the fractionation process: input, unbound and bound fractions. The intensities of the bands corresponding to representative PCR products amplified from the input, unbound and bound fractions are shown. (B) Genomic bisulfite sequencing profiles of *hTERT* promoter and proximal exonic region in wild-type HeLa and *MBD2* KD HeLa cells. A 550 bp region (positions nt –442 to nt +108) of the *hTERT* CpG island is presented on a schematic map. White dotted box, core promoter; black arrows, transcription start sites. Bisulfite-sequencing status of *hTERT* promoter and proximal exonic region is shown (number of analyzed clones for each cell line, $n = 10$). Each line represents a single-DNA template molecule. Filled and open squares, respectively, represent methylated and unmethylated CpGs. Sequencing was performed from two different regions that do not overlap (see Materials and Methods), leading to unanalyzed CpGs, barred squares. The slightly methylated region (positions nt –165 to nt –80) is framed. (C) MeCP2 and MBD1 do not compensate for MBD2 depletion at the *hTERT* promoter in *MBD2* KD HeLa cells. Representative examples of ChIP assays performed in *MBD2* KD HeLa cells are presented.

hypermethylated region of the *hTERT* core promoter but also covered all the hypermethylated *hTERT* CpG island (Figure 2A). As a control, results obtained for a previously identified MBD2-free hypermethylated promoter (28), *PARVG*, are also shown on Figure 2B. As expected, no MBD2-positive signal was observed along this hypermethylated promoter (Figure 2B).

The hypermethylated *hTERT* promoter remains free of MBD proteins in *MBD2*-depleted cells

In HeLa cells, among the MBD transcripts, *MBD2* mRNAs are the most abundant (27), suggesting that the selective binding of MBD2 to *hTERT* promoter might be due to its prominent expression. Thus, we investigated *hTERT* promoter occupancy in a HeLa clone cell line [*MBD2* knockdown (KD) HeLa cells] constitutively depleted in *MBD2* by a transgene expressing a siRNA targeting the mRNA coding for this protein (26).

Quantitative competitive RT-PCR assays indicated that 89–96% *MBD2* mRNA depletion was obtained and maintained over many passages in *MBD2* KD HeLa cells. This low level of *MBD2* transcripts was correlated with a very low level of MBD2 protein, which was

almost undetectable in western blot experiments (26). In addition, neither the amounts of MeCP2 and MBD1 transcripts nor the amounts of the corresponding proteins were altered by *MBD2* depletion (data not shown).

As MBD2 belongs to the DNA methylation machinery, a prolonged *MBD2* depletion might alter DNA methylation patterns. Therefore, the DNA methylation patterns of the *hTERT* promoter were determined in *MBD2* KD HeLa cells. DNA extracted from the cell lines was modified and amplified by PCR with primers specifically designed to amplify bisulfite-modified DNA sequence of the *hTERT* promoter and proximal exonic region. PCR fragments were cloned and sequenced. The analysis of 10 clones from wild-type HeLa cells and *MBD2* KD HeLa cells indicated that over the hypermethylated *hTERT* promoter and proximal exonic region, a small sequence (positions nt –165 to nt –80), corresponding to a part of the *hTERT* core promoter, was hypomethylated in the two cell lines analyzed (Figure 1B). Indeed, this region exhibit a low level of methylation (~30%) but no significant difference was observed between wild-type HeLa cells and HeLa cells depleted in *MBD2*. Thus, the methylation patterns of the *hTERT* promoter in *MBD2* KD HeLa cells were not altered by the absence of MBD2.

As expected, in HeLa cells depleted in *MBD2* proteins, ChIP assays indicated that MBD2 was no longer detected at the hypermethylated region of the *hTERT* core promoter (Figure 1C). Furthermore, this region seemed to remain free of MBD proteins since the immunoprecipitated fractions are depleted in the methylated *hTERT* promoter when anti-MeCP2 or anti-MBD1 antibodies were used in ChIP experiments (Figure 1C).

Thus, the hypermethylated region of the *hTERT* promoter is specifically targeted by MBD2 in HeLa cells and no redundancy between MBD2 and MeCP2 or MBD1 was observed at this hypermethylated region.

MBD2 depletion enhances *hTERT* gene transcription in HeLa cells

MBD2 is a member of the MBD protein family and their methylation-dependent repressive activities are now well established (29). Therefore, we investigated the potential involvement of MBD2 in the repression of the endogenous *hTERT* promoter in HeLa cells.

First, we determined the consequence of a transient *MBD2* depletion on the expression of *hTERT* in HeLa cells. At 24 h intervals after *MBD2*-specific siRNA transfection, RNA was extracted and the levels of *hTERT* and *MBD2* transcripts determined by RT-PCR assays were compared with their levels in HeLa cells transfected with a non-specific siRNA. Quantitative competitive RT-PCR assays (27) indicated that a reduction of ~90% in *MBD2* mRNA level was reached 24 h after transfection with *MBD2* siRNA and maintained >96 h after transfection (Figure 3A). Western blot analysis also showed a dramatic decrease in *MBD2* proteins in these cells (data not shown). In HeLa cells, 48 h after the *MBD2* siRNA transfection, an elevation of 1.2-fold of *hTERT* mRNA level was observed when compared with HeLa cells transfected with a non-specific siRNA. This stimulation increased to a maximum of 1.9 at 96 h after *MBD2* siRNA treatment. The expression level of *hTERT* was determined using relative RT-PCR (Figure 3A) and identical results were obtained with real-time RT-PCR for critical points (Figure 3B). It should be noted that neither the level of *MBD2* transcripts nor the level of *hTERT* transcripts were altered by non-specific siRNA transfection when compared with untreated wild-type HeLa cells (data not shown).

These data suggest that MBD2 actually represses *hTERT* expression in HeLa cells.

Stimulation of *hTERT* transcription in *MBD2* KD HeLa cells is reversed by ectopic expression of *Mbd2*

The *MBD2* KD HeLa cell line offers the opportunity to investigate the functional control of the specific repression of *hTERT* by MBD2. In this cell line, *MBD2* expression can be rescued using pRev-Mbd2, a vector coding for a mouse *Mbd2* RNA containing five silent point mutations and, therefore, refractory to siRNA-mediated decay (26). We also observed a 2-fold stimulation of *hTERT* expression in the *MBD2* KD HeLa cells as compared with wild-type HeLa cells (Figure 4). In these

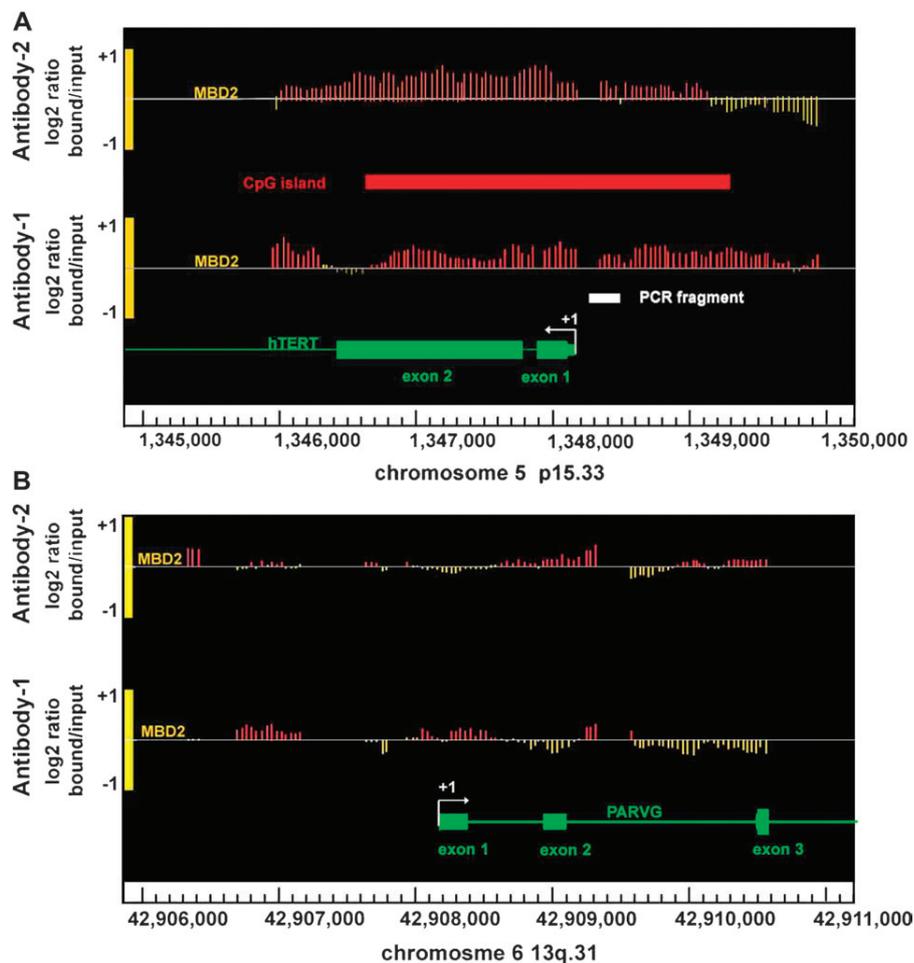


Fig. 2. ChIP-on-chip analysis of MBD2-binding sites on *hTERT* promoter. (A) Array peaks on *hTERT* promoter of MBD2 log₂ signal ratio (MBD2/input) values are shown below the Affymetrix' Integrated Genome Browser window. Gene is transcribed from right to left. *hTERT* CpG island is shown by a red box. DNA fragment analyzed by PCR following MBD2 ChIP is represented by a white box. (B) *PARVG* 5' end viewed as a MBD2-free control. Gene is transcribed from left to right.

cells, the ectopic expression of Mbd2 reduced *hTERT* transcription to a level similar to the level observed in wild-type HeLa cells (Figure 4), suggesting that MBD2 directly represses *hTERT* expression in HeLa cells. In wild-type HeLa cells, the abundance of MBD2 does not seem to be a limiting factor since the overexpression of MBD2 mediated by pRev-Mbd2 transfection did not modify *hTERT* expression level in cells containing normal amounts of MBD2 (Figure 4).

Our results demonstrate that MBD2 specifically binds to *hTERT* and represses its expression in HeLa cells.

MBD2 hTERT repression is specifically observed in hTERT-methylated telomerase-positive cancer cell lines and independent of the cellular types

Data obtained suggest that MBD2 might be a general repressor of *hTERT* transcription in *hTERT*-methylated telomerase-positive cells. In order to address this point, we tested whether the above-mentioned MBD2 KD-mediated *hTERT* stimulation in HeLa cells could be recapitulated in other *hTERT*-methylated cell lines. A functional study was undertaken in different cell lines exhibiting characteristic *hTERT* DNA methylation patterns: (i) MCF-7, LAN-1 and HepG2 cells, three telomerase-positive cancer cell lines showing an *hTERT* core promoter unmethylated or slightly methylated despite an hypermethylated CpG island (ii) NCCIT cells, an *hTERT*-unmethylated telomerase-positive teratocarcinoma cell line and (iii) MRC5 cells, an *hTERT*-unmethylated telomerase-negative normal embryonic cell

line. The *hTERT* core promoter methylation status of these cell lines and the corresponding transcript levels are shown in Figure 5. Control experiments indicated that these cell lines exhibited approximately the same level of MBD2 transcripts ($4.8 \pm 2.7 \times 10^6$ mRNA molecules per microgram of total RNA), with the exception of NCCIT (7×10^4 mRNA molecules per microgram of total RNA).

As observed in HeLa cells, a significant decrease of MBD2 transcripts (60–79%) was observed in the different cell lines 96 h after MBD2 siRNA transfection. Relative to the transfection with a non-specific siRNA, a ~2-fold increase in *hTERT* mRNA was observed in all *hTERT*-methylated cell lines, LAN-1, MCF7 and HepG2 cells. In contrast, MBD2 depletion did not affect *hTERT* expression in *hTERT*-unmethylated cell lines, NCCIT and MRC5 cells (Figure 6). Since MBD2 does not bind to unmethylated DNA (24,25), these data indicate that *hTERT* induction due to MBD2 depletion is not mediated by an indirect effect.

Taken together, the specific transcriptional repression of *hTERT* by MBD2 does not seem to be restricted to a particular cancer cell line since this effect was observed in cervix, breast, liver and neuroblastoma cancer cell lines. Thus, MBD2 seems to be a general repressor of *hTERT* in *hTERT*-methylated telomerase-positive cells.

Discussion

Most (~96%) of the CpG islands are unmethylated in normal cells (30), whereas hypermethylation of these sequences is a characteristic

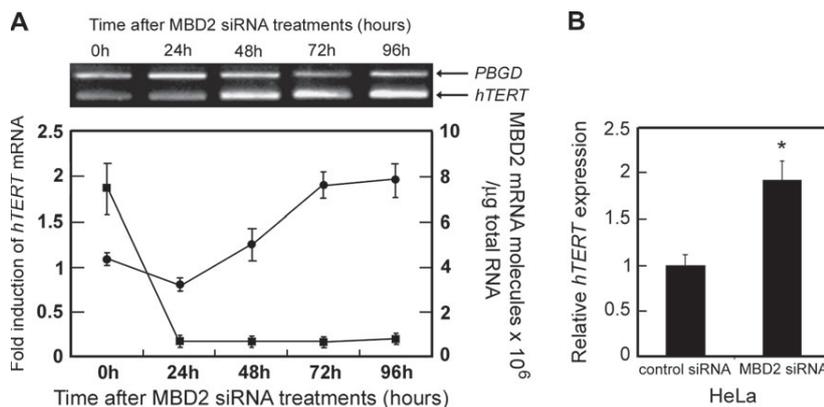


Fig. 3. Transient depletion of MBD2 proteins by MBD2-specific siRNA induces time-dependent stimulation of *hTERT* expression in HeLa cells. (A) HeLa cells were transiently transfected, with either MBD2 siRNA or a negative control scrambled siRNA. After 24, 48, 72 or 96 h following the transfection, RNA was extracted and the efficiency of the MBD2 siRNA treatment was determined by quantification of *MBD2* mRNA as described previously (27). *hTERT* mRNA levels were monitored by relative RT-PCR. The fold induction of *hTERT* expression was calculated from the ratio *hTERT* mRNA/*PBGD* mRNA in MBD2 siRNA-transfected cells versus scrambled siRNA-transfected cells. Mean values (± standard deviation) obtained from at least three independent transfection experiments are shown. Filled squares, *MBD2* mRNA level; filled circles, fold induction of *hTERT* mRNA. (B) Quantitative RT-PCR expression analysis of *hTERT* in HeLa cells 96 h after transfection with either a negative control siRNA or MBD2 siRNA. Real-time RT-PCR was done on cells lines and *hTERT* expression was normalized to β -actin. The relative level of each mRNA was calculated on the basis of the two standard curves using the relative quantification method. At least three independent determinations of fold differences were used to calculate the average fold difference values and associated standard deviation ($P = 0.0027$; *t*-test).

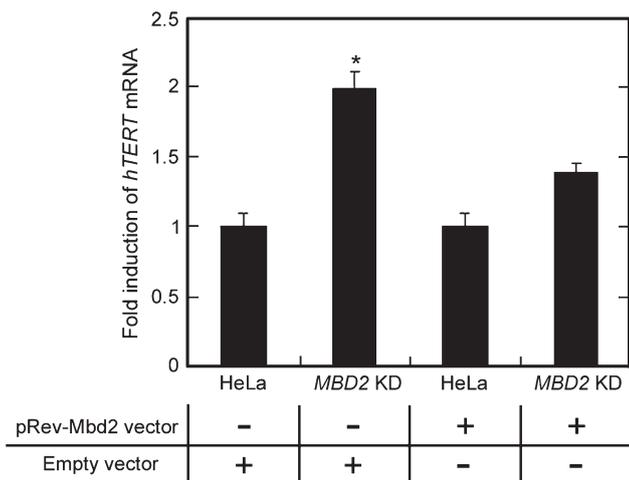


Fig. 4. Mbd2 expression rescues the reduction of *hTERT* transcript in *MBD2* KD HeLa cells. The transcriptional expression of *hTERT* was analyzed by quantitative RT-PCR 48 h after lipofectamine transfection of HeLa and *MBD2* KD HeLa cells using pRev-Mbd2, an Mbd2 vector expressing a transcript resistant to RNAi or an empty pGL3 basic vector. The relative *hTERT* mRNA in pRev-Mbd2-transfected cells or in pGL3-transfected cells or in KD HeLa cells was calculated using the same reference: the *hTERT* mRNA level in untransfected HeLa cells. Mean values obtained from at least three independent transfection experiments are shown ($P = 0.0035$, *t*-test).

of cancer cells (1). These aberrant DNA methylation patterns have been correlated with the transcriptional silencing of genes undergoing such alterations at their 5' end CpG islands. The *hTERT* gene has provided an interesting exception since a bimodal alteration of the DNA methylation status of its 5' end CpG island is associated with its expression in ~85% of cancer cells and tumors tissues (4,5). The large CpG island, 2.6 kb in length, lying from nt -1102 to nt +1519 from the *hTERT* transcription, is hypermethylated at the exception of a short region (positions nt -165 to nt -80) that is unmethylated or slightly methylated despite highly methylated border regions (16,17). This particular pattern of methylation seems crucial for establishing

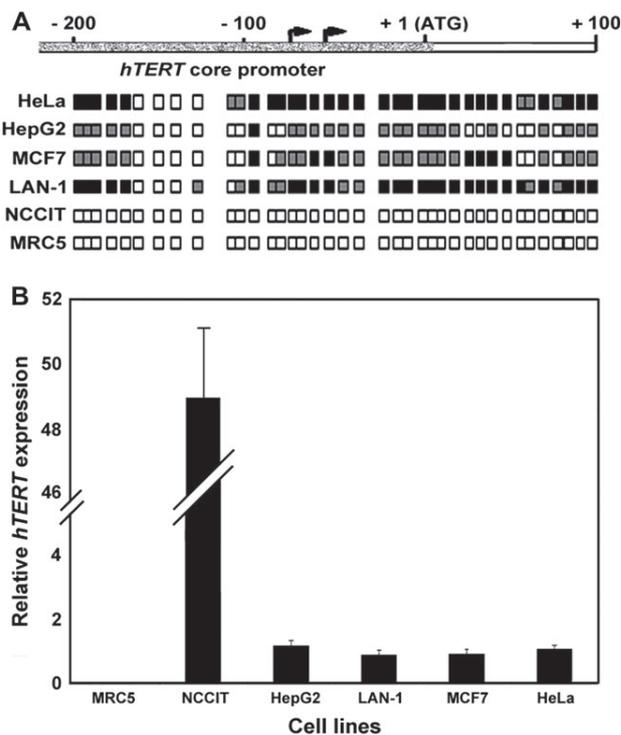
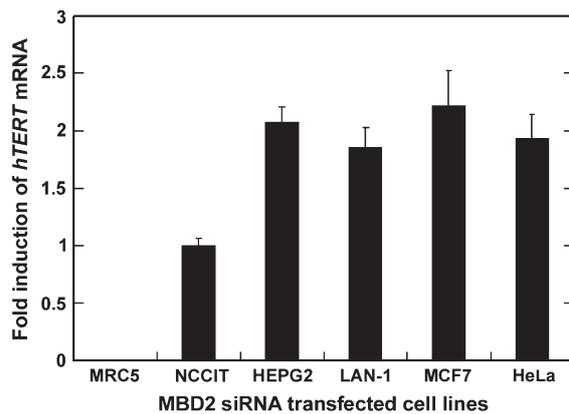


Fig. 5. DNA methylation patterns of *hTERT* and the corresponding transcript levels in different cell lines. (A) Bisulfite sequencing of five telomerase-positive cancer cell lines (HeLa, MCF7, HepG2, LAN-1 and NCCIT) and one normal telomerase-negative fibroblast cell line (MRC5). Bisulfite-treated DNA was PCR amplified using a primer set spanning the end of the promoter and the proximal exonic region of *hTERT*. PCR products were then directly sequenced. The analyzed 314 bp region (-206 to +108) is presented on a schematic map and results are shown for each cell lines. Filled and open squares, respectively, represent complete methylated and unmethylated CpG sites, whereas gray squares correspond to partial methylated CpG sites. (B) Quantitative expression of *hTERT* in MRC5, NCCIT, HepG2, LAN-1, MCF7 and HeLa cells. Real-time RT-PCR was done on cell lines, and β -actin was used as a reference.



Efficiency of MBD2 siRNA	91% ± 9%	63% ± 5%	63% ± 7%	73% ± 6%	72% ± 3%	87% ± 7%
--------------------------	----------	----------	----------	----------	----------	----------

Fig. 6. Depletion of MBD2 by siRNA results in transcriptional activation of *hTERT* in *hTERT*-methylated telomerase-positive cancer cell lines. Fold change of *hTERT* expression between MBD2-depleted cells (MBD2 siRNA-transfected cells) and control cells (scrambled siRNA-transfected cells). *hTERT* expression was monitored by quantitative RT-PCR 96 h after transfection. Mean values (± standard deviation) obtained from at least three independent transfection experiments are shown.

hTERT expression at a basal level. Indeed, hypermethylation of CpG islands seems to be a crucial event in carcinogenesis. Thus, the methylation-free region in *hTERT* promoter may result from antagonistic pressure between the mechanisms leading to aberrant methylation and the need to keep *hTERT* expressed for unlimited life span of cancer cells.

A body of evidence has been accumulated concerning association between hypermethylation of CpG islands, transcriptional silencing and MBD-proteins binding (21). ChIP experiments and ChIP-on-chip analysis indicate that MBD2 associated the hypermethylated CpG island of *hTERT*. In this context, MBD2 seems to be a limiting factor rather than a transcriptional silencer.

Recently, a large screening of tumor suppressor gene promoters in 10 cell lines showed that human cancer cell lines tend to use a particular MBD protein (31). Furthermore, in MRC5 cells, ChIP experiments have indicated that MeCP2 and MBD2 proteins have non-overlapping binding specificities *in vivo* (23). These data, as well as our data are in favor of the ‘one gene–one MBD’ hypothesis, at least for some genes. Indeed, MBD2 was specifically associated with the methylated region of the *hTERT* CpG island, whereas MeCP2 and MBD1 were not detected at this locus. Moreover, MBD2 depletion did not induce MeCP2 or MBD1 binding at this methylated area in HeLa cells. Thus, the specificity of MBD proteins does not seem to be driven by their relative concentrations in a cell line, as it was suggested in another study (31): a strong expression of a particular MBD is not necessarily associated with its preferential use in promoters.

Several studies have shown that transcriptional activation could be realized upon depletion of MBD proteins by RNA interference (RNAi) (26,28). A large-scale microarray analysis indicated that 15% of 6386 genes analyzed exhibit an increased expression change between untreated and triple MBD-depleted cells (28). It should be noted that for single-MBD interference, MBD2 depletion was the protein most commonly involved in the observed release of gene silencing by far (28). These data suggest that MBD2 plays an important role in methylation-dependent gene silencing. MBD2 depletion mediated by RNAi stimulates *hTERT* expression, either in stable KD HeLa clones or in transiently transfected HeLa cells. A 2-fold stimulation of *hTERT* expression upon MBD2 depletion was observed, suggesting that MBD2 plays an important role in the regulation of this gene. Indeed, in human cell lines, fold changes induced by MBD2

depletion is also of the same range, 2-fold for NBR2 induced by MBD2 siRNA (26) and microarray analysis of MBD2-depleted cells exhibited a mean fold change of 7 (28). Furthermore, the fold change observed in Mbd2-deficient mouse cells is not very different from our own data. For example, it has been shown (32) that Mbd2^{-/-} fibroblasts had 3-fold higher levels of Xist than wild-type cells. Moreover, elevated *hTERT* transcription upon MBD2 depletion is not limited to cervix (HeLa cells) since this effect was observed in breast, liver and neuroblastoma cancer cell lines. MBD2 RNAi experiments in cell lines exhibiting unmethylated *hTERT* CpG island (MRC5 and NCCIT) did not affect *hTERT* transcription indicating that MBD2 specifically and directly represses *hTERT* expression in methylation-dependent manner. Collectively, these data strongly suggest that MBD2 is a general repressor of *hTERT* expression in cancer cells.

A considerable number of transcription factors have been proposed as negative regulators of *hTERT* transcription. Nevertheless, direct evidence indicating that *hTERT* gene transcription is downregulated remains to be firmly established. Thus, MBD2 may represent a new factor directly involved in the negative regulation of *hTERT* expression. In cancer cells, MBD2 seems to play a specific role since its binding to the hypermethylated part of the *hTERT* promoter establishes a direct link between a common DNA alteration of tumor cells, CpG hypermethylation and a reduced level of *hTERT* transcription. The binding of MBD2 represents an additional layer for the control of *hTERT* expression, thereby underlying the importance of *hTERT* regulation in immortalized and cancers cells.

Supplementary material

Supplementary Figure S1 and Table S1 can be found at <http://carcin.oxfordjournals.org/>

Funding

Ligue Nationale contre le Cancer (Comité du Rhône and Comité de la Loire); Institut National contre le Cancer (EpiPro, CircBio, 419807-37/NG-LC); Swiss National Science Foundation (3100A0-101732, 3100A0-113505).

Acknowledgements

Conflict of Interest Statement: None declared.

References

- Ehrlich,M. (2002) DNA methylation in cancer: too much, but also too little. *Oncogene*, **21**, 5400–5413.
- Esteller,M. (2007) Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat. Rev. Genet.*, **8**, 286–298.
- Herman,J.G. *et al.* (2003) Gene silencing in cancer in association with promoter hypermethylation. *N. Engl. J. Med.*, **349**, 2042–2054.
- Kim,N.W. *et al.* (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science*, **266**, 2011–2015.
- Stampfer,M.R. *et al.* (2003) Human epithelial cell immortalization as a step in carcinogenesis. *Cancer Lett.*, **194**, 199–208.
- Yasumoto,S. *et al.* (1996) Telomerase activity in normal human epithelial cells. *Oncogene*, **13**, 433–439.
- Goueli,B.S. *et al.* (2003) Regulation of telomerase reverse transcriptase gene activity by upstream stimulatory factor. *Oncogene*, **22**, 8042–8047.
- Poole,J.C. *et al.* (2001) Activity, function, and gene regulation of the catalytic subunit of telomerase (hTERT). *Gene*, **269**, 1–12.
- Renaud,S. *et al.* (2005) CTCF binds the proximal exonic region of hTERT and inhibits its transcription. *Nucleic Acids Res.*, **33**, 6850–6860.
- Dessain,S.K. *et al.* (2000) Methylation of the human telomerase gene CpG island. *Cancer Res.*, **60**, 537–541.
- Devereux,T.R. *et al.* (1999) DNA methylation analysis of the promoter region of the human telomerase reverse transcriptase (hTERT) gene. *Cancer Res.*, **59**, 6087–6090.
- Guilleret,I. *et al.* (2002) Hypermethylation of the human telomerase catalytic subunit (hTERT) gene correlates with telomerase activity. *Int. J. Cancer*, **101**, 335–341.

13. Nomoto, K. *et al.* (2002) Methylation status and expression of human telomerase reverse transcriptase mRNA in relation to hypermethylation of the p16 gene in colorectal cancers as analyzed by bisulfite PCR-SSCP. *Jpn. J. Clin. Oncol.*, **32**, 3–8.
14. Renaud, S. *et al.* (2007) Dual role of DNA methylation inside and outside of CTCF-binding regions in the transcriptional regulation of the telomerase hTERT gene. *Nucleic Acids Res.*, **35**, 1245–1256.
15. Horikawa, I. *et al.* (1999) Cloning and characterization of the promoter region of human telomerase reverse transcriptase gene. *Cancer Res.*, **59**, 826–830.
16. Zinn, R.L. *et al.* (2007) hTERT is expressed in cancer cell lines despite promoter DNA methylation by preservation of unmethylated DNA and active chromatin around the transcription start site. *Cancer Res.*, **67**, 194–201.
17. Guilleret, I. *et al.* (2004) Unusual distribution of DNA methylation within the hTERT CpG island in tissues and cell lines. *Biochem. Biophys. Res. Commun.*, **325**, 1037–1043.
18. Yi, X. *et al.* (2001) Quantitation of telomerase components and hTERT mRNA splicing patterns in immortal human cells. *Nucleic Acids Res.*, **29**, 4818–4825.
19. Ducrest, A.L. *et al.* (2001) Regulation of human telomerase activity: repression by normal chromosome 3 abolishes nuclear telomerase reverse transcriptase transcripts but does not affect c-Myc activity. *Cancer Res.*, **61**, 7594–7602.
20. Hendrich, B. *et al.* (1999) The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature*, **401**, 301–304.
21. Bird, A.P. *et al.* (1999) Methylation-induced repression—belts, braces, and chromatin. *Cell*, **99**, 451–454.
22. Fatemi, M. *et al.* (2006) MBD family proteins: reading the epigenetic code. *J. Cell Sci.*, **119**, 3033–3037.
23. Klose, R.J. *et al.* (2005) DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. *Mol. Cell*, **19**, 667–678.
24. Ng, H.H. *et al.* (1999) MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat. Genet.*, **23**, 58–61.
25. Wade, P.A. *et al.* (1999) Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nat. Genet.*, **23**, 62–66.
26. Auriol, E. *et al.* (2005) Specific binding of the methyl binding domain protein 2 at the BRCA1-NBR2 locus. *Nucleic Acids Res.*, **33**, 4243–4254.
27. Billard, L.M. *et al.* (2002) MeCP2 and MBD2 expression during normal and pathological growth of the human mammary gland. *Oncogene*, **21**, 2704–2712.
28. Lopez-Serra, L. *et al.* (2008) Unmasking of epigenetically silenced candidate tumor suppressor genes by removal of methyl-CpG-binding domain proteins. *Oncogene*, **27**, 3556–3566.
29. Klose, R.J. *et al.* (2006) Genomic DNA methylation: the mark and its mediators. *Trends Biochem. Sci.*, **31**, 89–97.
30. Weber, M. *et al.* (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.*, **39**, 457–466.
31. Lopez-Serra, L. *et al.* (2006) A profile of methyl-CpG binding domain protein occupancy of hypermethylated promoter CpG islands of tumor suppressor genes in human cancer. *Cancer Res.*, **66**, 8342–8346.
32. Barr, H. *et al.* (2007) Mbd2 contributes to DNA methylation-directed repression of the Xist gene. *Mol. Cell. Biol.*, **27**, 3750–3757.

Received August 20, 2008; revised October 3, 2008;
accepted October 11, 2008

Annexe 2

Liste des publications

- (1) **Perriaud L.**, Morel A.P., Lachuer J., Dante R., Specific Distribution of the Methyl-CpG Binding Domain Protein 2 at methylated DNA sequences during tumoral progression, in a syngenic model of breast cancer. *En cours de redaction*
- (2) Chatagnon, A., **Perriaud L.**, Nazaret N., Croze S., Benhattar J., Lachuer J., Dante R., Methyl-CpG Binding Domain Protein 2, a DNA Methylation-Dependent Gene Silencer, Genome-Wide Analysis. *Soumis à FASEB J.*
- (3) **Perriaud L.**, Lachuer J., Dante R., Relationship Between the CpG Content of Gene Promoter and Transcriptional Repression by the Methyl-CpG Binding Domain Protein 2 in Normal Human Fibroblastic Cells. *Soumis à PloS One.*
- (4) Chatagnon, A., Bougel, S., **Perriaud, L.**, Lachuer, J., Benhattar, J., Dante, R., Specific association between the methyl-CpG binding domain protein 2 and the hypermethylated region of the human telomerase reverse transcriptase promoter in cancer cells. *Carcinogenesis*, 2009. 30(1): 28-34.

Communications affichées

Chromatin Plasticity Network & Epigenome Network of Excellence - Joint Conference: Advances in Epigenetics, 10-14 Mai 2010, San Feliu de Guixols, Espagne.

Perriaud, L., Croze, S., Morel, A.-P., Lachuer, J., Dante, R., *Genome-wide analysis of epigenetic events during tumor progression in a syngenic model of breast cancer.*